

DEX-0113

BL

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : C12N 15/12, 1/21, 5/10, C07K 14/47, 16/18, A61K 38/17, G01N 33/50, C12Q 1/68	A2	(11) International Publication Number: WO 00/11168 (43) International Publication Date: 2 March 2000 (02.03.00)
(21) International Application Number: PCT/US99/19052 (22) International Filing Date: 20 August 1999 (20.08.99) (30) Priority Data: 09/138,132 21 August 1998 (21.08.98) US (63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application US 09/138,132 (CIP) Filed on 21 August 1998 (21.08.98) (71) Applicant (for all designated States except US): PRINCETON UNIVERSITY [US/US]; New South Building, 5th floor, P.O. Box 36, Princeton, NJ 08544-0036 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): LEMISCHKA, Ihor [US/US]; 4 Firestone Court, Princeton, NJ 08540 (US). MOORE, Kateri [US/US]; 248 Hawthorne Street, Princeton, NJ 08540 (US). (74) Agents: YAMIN, Michael, A. et al.; Klauber & Jackson, 411 Hackensack Avenue, Hackensack, NJ 07601 (US).	(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>Without international search report and to be republished upon receipt of that report.</i>	
(54) Title: GENES THAT REGULATE HEMATOPOIETIC BLOOD FORMING STEM CELLS AND USES THEREOF (57) Abstract <p>The present invention provides an isolated nucleic acid derived from an isolated hematopoietic stem cell. The present invention additionally provides an isolated hematopoietic stem cell specific protein or a portion thereof encoded by the provided nucleic acid. In addition, the present invention provides a method for generating a stem cell/progenitor cell from a primitive hematopoietic cell. Also, the present invention further provides a method for identifying the presence in a sample of a compound that modulates hematopoietic stem cell activity. The present invention additionally provides a molecularly defined primitive hematopoietic stem cell. Finally, the present invention provides a method of ex vivo expansion of hematopoietic stem cells.</p> <p style="text-align: center; font-weight: bold; font-size: 1.2em;">BEST AVAILABLE COPY</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

GENES THAT REGULATE HEMATOPOIETIC BLOOD FORMING STEM CELLS AND USES THEREOF

5

GOVERNMENTAL SUPPORT

This invention was made with the support of National Institutes of Health Grant No. RO1 CA45339-09 and American Cancer Society No. DHP-144/01. The United States government may have certain rights to this invention.

Throughout this application, various publications are referenced by number. Full citations for these publications may be found listed at the end of the specification and preceding the Claims. The disclosures of these publications in their entireties are hereby incorporated by reference into this application in order to more fully describe the state of the art. A Sequence Listing is provided.

15

FIELD OF THE INVENTION

The present invention relates to hematopoietic stem cells and the stem cell and support cell genes that support stem cell replication and differentiation.

20

BACKGROUND OF THE INVENTION

The adult hematopoietic system is organized as a hierarchy of cells with decreasing self-renewal and multilineage differentiation potential. This is accompanied by progressively larger numbers of more mature cells and an increasing tendency to be in active cell cycle (Lemischka, I.R., 1992; Morrison, S.J., et al. 1995). Collectively, the properties of this hierarchical system result in the balanced, lifelong production of at least eight distinct cell lineages. A population of stem cells establishes the entire hierarchy; therefore, in order to understand the fundamental mechanisms responsible for normal hematopoiesis it is ultimately necessary to understand the biology of the stem cells.

25

Most information concerning the biology of stem cells has been obtained from the mouse model. In this system, the most critical, characteristic property of the stem cell population has been defined; that is, its ability to reconstitute a normal blood system in a transplanted host. A number of variations on the basic transplantation assay have been

30

described (Harrison, D.E. 1980; Spangrude, G.J., et al. 1995). All of these systems, together with the appropriate donor vs. host or clonotypic markers have rigorously defined the most primitive stem cells and have provided a description of their developmental properties. Perhaps the most striking characteristics of this cell population come from retroviral "marking" studies (Leminskka, I.R. 1992). These studies clearly show that a single stem cell clone is both necessary and sufficient, not only to sustain lifelong, multilineage hematopoiesis in one primary recipient but in numerous secondary animals. This illustrates the remarkable proliferative potential of the stem cell and directly demonstrates stem cell self-renewal. A major advance in mouse stem cell biology was the development of strategies which facilitate the substantial enrichment of stem cell activity (Bauman, J.G., et al. 1988; Spangrude, G.J., et al. 1988; Jordan C.T., et al 1990). Purification procedures enabled the first direct approaches to unravel the mechanisms responsible for the unique biological properties of the stem cell population. A key observation was that the phenotypically defined stem/progenitor cell population is heterogeneous with respect to in vivo functional properties (Fleming, W.G., et al 1993; Li, C.L. and Johnson G.R. 1992; Spangrude, G.J. and Johnson, G.R. 1990, Jones R., et al. 1990; Uchida, N., et al. 1993). In addition to the in vivo repopulating cells, other primitive progenitor cells are often contained in a purified population (Weilbaecher, K., 1991; Trevisan, M. and Iscove, N.N. 1995; Ogawa, M. 1993). These can be assayed in a variety of in vitro systems. Whether all of these in vivo and in vitro activities represent discrete cell subpopulations or whether there is a continuum of functional potential is still an unanswered question. Recent studies have suggested distinct physical properties for functionally different activities within the primitive population Morrison, S.J., and Weissman, I.L. 1994); Morrison, S.J. et al. 1997; Jones, R.M., et al. 1996). A second set of observations revealed an inverse correlation between a tendency for active cell cycling and primitive, uncommitted developmental potential in BM (Spangrude, G.J., and Johnson, G.R. 1990). In fetal liver a higher proportion of primitive stem cells is actively cycling (Fleming, W.H., et al. 1993). Moreover, it has been shown that fetal stem cells are more potent than adult stem cells in LTRA (Jordan, C.T., et al. 1995; Pawliuk, R., et al 1996). These are exciting observations because they suggest that rapid stem cell cycling can be compatible with the maintenance of primitive in vivo activity. Very recent studies suggest that the adult BM stem cell compartment may in fact be cycling at a very slow rate (Bradford, G.B., et al. 1997). Clearly, stem cell cycle regulation is a critical

area for investigation. Stem cell purification has facilitated studies aimed at ex vivo maintenance or expansion of the most primitive, transplantable stem cell. Most culture systems strongly favor a differentiation process (Van der Sluijs, J.P., et al. 1993; Traycoff, C.M., et al. 1996; Peters, S.O., et al. 1995; Knobel, K.M., et al. 1994).
5 However, several recent reports have been encouraging. It has been shown that colonies grown in defined cytokines can retain not only myeloid and erythroid but also lymphoid potentials (Ball, T.C., et al. 1995). Moreover, the short-term (2-3 weeks) maintenance of LTRA has been demonstrated in suspension cultures supported by IL6, IL11, together with ckit ligand (KL) or flk2/flt3 ligand (FL) (Yonemura, Y.H., et al. 1997). A recent
10 report has shown that colonies initiated in cytokine-supplemented semisolid cultures retain LTRA (Trevisan, M., et al. 1996). The studies described herein have developed a stromal cell line supported system which quantitatively maintains LTRA for an extended (4-7 weeks) time (Moore, F.A., et al. 1997).

In the human system it is clearly not possible to do the same kind of extensive
15 marking and transplantation assays. However, several xenograft model systems have been developed to assess the in vivo behavior of human stem cells (Traycoff, C., et al. 1994; Turner, C., et al. 1996; Cashman, J., et al. 1997). Some of these experiments can be done quantitatively in limiting dilution (Bhatia, M., et al. 1997). A very recent study has demonstrated a common proviral integration site in granulocyte macrophages and T-
20 cells derived from beige/nude/XID mice 7-11 months after engraftment with genetically transduced human stem cells (Nolta, J.A., et al. 1996). This important study paves the way for precise in vivo clonal analyses. The largest amount of functional information about human stem/progenitor cells has been obtained in vitro using a wide range of stromal cell and cytokine supported culture systems. It is not possible herein to describe
25 and properly accredit all of the important studies, however several advances deserve mention. The long-term culture-initiating cell (LTCIC) assay measures the in vitro production of colony forming cells (CFC) after periods of at least five weeks in culture (Sutherland, H.S., et al. 1989). The cells producing these CFC derive from a population of cells which, at least to some extent, probably overlaps with the most primitive
30 compartment. The maintenance and expansion of primitive functional abilities in this culture system has recently been documented (Petzer, A.L., et al. 1996). A variation on this assay system, the extended LTCIC (ELTCIC) has been suggested to measure an even more primitive cell population in BM and CB (Hao, Q.L., et al. 1996). A very exciting

prospect for the near future will be the integration of the various in vivo xenograft assays with the in vitro LTCIC and ELTCIC systems. Some very recent efforts have suggested that the NODSCID xenograft system and the LTCIC assay may measure distinct stem/progenitor subsets (Larochelle, A.J., et al. 1997). Clearly, much more work needs to be done, however, it may be anticipated that the ELTCIC system will provide the "bridge" in this continuum. Collectively, and including the various strictly cytokine-driven systems, the above studies illustrate the current possibility to accurately and quantitatively reveal the majority (if not all) functional entities in the human stem/progenitor cell hierarchy. The physical characterization and purification of human stem/progenitor cells has proceeded along lines which are parallel to the mouse system. Indeed, because of clinical impetus, it can be argued that they are further advanced. Thus, as measured in the range of assays discussed above, the consensus physical phenotype of the most primitive portion of the human stem/progenitor hierarchy is CD34+Lin-CD38- (Terstappen, L.W.M.M., et al. 1991). The CD34+Lin-CD38+ subset contains less primitive, more committed cellular entities. Other studies have shown that, similar to the mouse, low level expression of Thyl (CD90) is a feature of the primitive human stem cell (CD34+Lin-CD90+)(Baum, C.M., et al. 1992; Craig, W., et al. 1993). Most CD90+ cells in this compartment are CD38-. Therefore, the consensus phenotype can be described as CD34+Lin-CD90+ (Craig, W., et al. 1993). Two potential differences with the murine system can now be highlighted. First, a very recent and elegant study has shown that the most primitive mouse stem cell may in fact be CD34-/lo (Osawa, M., et al. 1996). Whether this is a genuine difference or whether it reflects the ability to perform more accurate long-term engraftment studies in the mouse remains to be determined. Second, it has been suggested that in the mouse, CD38 expression is a positive indicator for primitive stem cell function in a purified population (Randall, T.D., et al. 1996). As in the mouse, human stem/progenitor cells have been identified and purified from various sources. These include: adult BM (Baum, C.M., et al. 1992), CB (DiGiusto, D.L., et al 1996), fetal liver (Craig, W., et al. 1993) and peripheral blood stem cells after various mobilization protocols (Murray, L., et al. 1994). Similar to the data obtained in the murine system, comparative studies reveal that, in general, the basic and fundamental functional properties of stem/progenitor cells are shared regardless of the tissue source. There are however, significant functional and physical differences.

Numerous insights into hematopoietic molecular control mechanisms have come from gene-targeting studies in mice. Mutations in specific genes, most notably, those encoding transcription or DNA-binding proteins, have profound cell-intrinsic, global or lineage-specific effects on hematopoietic development (Shivdasani, R.A., and Orkin, S.H. 1996; Orkin, S.H. 1996). In the latter cases, it is tempting to speculate that the phenotypes result from defects in the commitment process. However, malfunctions in the commitment decision to "set up" a program of differentiation are difficult to distinguish from malfunctions in the differentiation program itself. Two gene products, AML1 (CBF2) and SCL (tal-1) appear to be necessary global regulators of hematopoiesis (Wang, W., et al. 1996; Okuda, T., et al. 1996; Porcher, C., et al. 1996; Robb, L., et al. 1996). Whether these molecules act to specify a hematopoietic stem cell or by other means is an open question. Interestingly, both of these molecules play roles in leukemic transformation. A very important gain of function study documents the apparent ability of HOXB4 to increase primitive cell numbers without significant impairment of differentiation abilities (Sauvageau, G., et al. 1995). Together with observations that HOXA9 is translocated in myeloid leukemia (Nakamura, T., et al. 1996; Borrow, J., et al. 1996), these studies suggest an important hematopoietic role for homeobox proteins. Without question, the above and other studies have identified important regulators of hematopoiesis. However, in almost all cases these regulators were first identified in other systems. The opposite approach is to directly search for stem cell regulators in stem cells, Graf, L., and Torok-Storb, B. 1995; Yang, Y., et al. 1996). The present invention solves these problems.

SUMMARY OF THE INVENTION

The human hematopoietic stem/progenitor cell population has been extensively characterized according to physical and antigenic criteria as well as in a variety of in vitro and in vivo assay systems. Collectively the human studies have revealed similarities to the hierarchical stem/progenitor cell organization defined in the murine system. In spite of significant strides in the identification of cytokines which can act on stem cells, it has not been possible to define a system where undifferentiated expansion of the most primitive stem cell population occurs. Similarly, it has not been possible to direct differentiation along lineage-specific pathways. These limitations, which also apply to the murine system, have hampered the elucidation of regulatory mechanisms which

mediate the most fundamental aspect of stem cell biology; that is, the decision to self-renew or commit to differentiation. As a consequence, very little is known about the molecular biology of the most primitive hematopoietic stem cell in any organism. It was hypothesized that the regulation of primitive stem cells will be mediated at least in part by the products of genes which are uniquely or predominantly expressed in these cells. One precedent for an important, differentially-expressed molecule is the flk2/flt3 receptor tyrosine kinase. There presumably are other important and differentially expressed gene products. Therefore, it is an object of the present invention to identify these molecules and address their functional roles. Specifically, an aspect of identification of gene expression patterns specific to primitive human stem cells is the molecular phenotype of the human stem cell. The present invention describes methods to define the profile of genes specifically expressed in undifferentiated human stem/progenitor cell populations.

A primary focus of the present invention is on primitive cells isolated from normal bone marrow (BM) samples. The present invention further comprehends use of other sources of stem cells, such as umbilical cord blood (CB).

The methods of the present invention combine diverse technical approaches and sophisticated bioinformatic analyses.

This invention further provides methods to identify genes whose expression can be modulated by cytokine or stromal-dependent culture and/or by cell-cycle status.

Another object of the present invention is to provide methods for the functional characterization of human stem cell-specific gene products. An aspect of this invention is a method to facilitate the functional characterization of specifically expressed gene products as candidate regulators of a variety of stem/progenitor cell processes. In particular, a provided method uses an in vitro system which approximates many characteristic properties of normal stem cells to analyze positive and negative regulation of proliferation, cell-cycle parameters, apoptosis and commitment.

It is a further goal of the present invention to provide a necessary (and usually missing) component for stem cell gene-expression screens; that is, the ability to quickly assess the function of extensive panels of genes.

It is also an object of the present invention to provide a method for the functional identification of stem cell regulators. An aspect of this invention is a facile screening method for "categorizing" large populations of specifically-expressed molecules according to their potential roles in a variety of stem/progenitor cell processes. Gain of

function as well as loss of function approaches are contemplated by the present invention. This method does not rely on any prior knowledge of nucleotide or predicted protein sequence.

Accordingly, the present invention provides an isolated nucleic acid
5 derived from an isolated hematopoietic stem cell, the isolated nucleic acid comprising the following characteristics: (1) specifically expressed in the hematopoietic stem cell; and (2) encoding a hematopoietic stem cell – specific protein.

The present invention additionally provides an isolated hematopoietic stem cell
10 specific protein or a portion thereof encoded by the provided nucleic acid. The present invention further still provides a nucleic acid probe capable of specifically hybridizing with the provided nucleic acid under standard hybridization conditions.

Also, the present invention provides an antibody capable of specifically binding to the provided protein without substantially cross-reacting with a non-stem cell specific
15 protein or homologs thereof under conditions permissive to antibody binding. Additionally, the present invention provides a cell capable of producing the provided antibody.

In addition, the present invention provides a method for identifying the presence of a primitive hematopoietic stem cell in a sample comprising nucleic acids specifically
20 expressed in hematopoietic stem cells. Further still, the present invention provides a method for generating a stem cell/progenitor cell from a primitive hematopoietic cell in a sample.

The present invention further provides a method for identifying the presence in a sample of a compound that modulates hematopoietic stem cell activity.

25 The present invention even further still provides a method for identifying primitive hematopoietic stem cell-specific nucleic acids.

Also the present invention additionally provides a molecularly defined primitive hematopoietic stem cell.

Yet additionally, the present invention provides a method for treating a condition
30 in a subject comprising administering to the subject a therapeutically effective amount of a provided pharmaceutical composition. The present invention provides a method of introducing an exogenous nucleic acid into a hematopoietic stem cell.

Finally, the present invention provides a method of ex vivo expansion of hematopoietic stem cells. The expanded cells are available to receive exogenous genes, including by retroviral or other vectors which require a round of replication. Alternatively, the cells are available for transplantation either autologously or heterologously.

5

BRIEF DESCRIPTION OF THE FIGURES

FIGURE 1. Schematic "flow of information" strategy.

10

FIGURE 2. Mammalian stem cell system (black circle = strong hybridization signal, gray circle = detectable hybridization signal, and open circle = no detectable hybridization signal). The Smc-34 cDNA is a completely novel sequence with a predicted leucine zipper and several potential membrane spanning domains.

15

FIGURES 3A-3B. (right panel), the control, non-subtracted RDA cDNA population (38-) contains b-actin sequences which are missing in the two subtracted RDA populations, 38- and 38-)38+. A differentially expressed gene (HDD-2, described below) is enriched in the 38- RDA population and at least retained in the 38-)38+ RDA population (Figure 3A, left panel). Two, bi-directional, RDA cDNA populations (38- and the converse 34-)38) were used to probe (See Figure 3) duplicate arrays of a subtracted 38- library (Figure 3B).

20

FIGURE 4A-4E. **Figure 4A:** 34B4 (SEQ.ID.No.: 69) is closely related to a gene encoding TINUR. The sequence homologies and restricted expression pattern of 34A5 is shown in **Figures 4B and 4C**. In **Figure 4C and 4D** (and also 5a, 5b, and 6b) there are twenty-one samples of capfinder-amplified cDNA from various hematopoietic populations. From left to right these are: four CD34+Lin- populations, three CD34+Lin-CD90+ populations, two CD34- populations, four CD34+Lin-CD38+ (obtained from the same BMs as the CD38- samples in lanes 1-4), two CD34+Lin- samples (obtained from the same BMs as the CD90+ samples in lanes 5 and 6), three CD34+Lin- populations obtained after 1, 2 or 4 days of culture and finally their three CD34+Lin-CD38+ counterparts. The 38G2 cDNA (SEQ.ID.No.: 70) is closely homologous to the

30

LTG9/MLLT3 gene located on 9q22 and involved in t(9;11) leukemia (Iida, S., et al. 1993) (Figure 4E).

FIGURES 5A-5B. G0S3, a fos-related gene (Heximer, S., et al. 1996) (Figure 5A) and
5 (2) HLA-DR (Figure 5B).

FIGURES 6A-6D. Figure 6A. HDD-2 is about 500 bp. It contains a short open reading frame of 89 amino acids (SEQ.ID.No.: 71). The predicted peptide sequence is shown. Figure 6B. The expression profile of HDD-2 demonstrates that it is stem cell restricted.
10 Figure 6C. HDD-2 hybridization to a dot blot with numerous human pA+ mRNA samples (Clontech). HDD-2 hybridization is only visible in kidney (the other "spots" are background). Figure 6D. Genomic Southern blot confirmed that HDD-2 corresponds to a single-copy human gene.

15 **FIGURE 7.** Immunoprecipitation analysis of protein extracts using rat IgG2b isotype control antibody (IgG) or AA4.1 mAb. Protein extracts were prepared from D2N cells; M2.4 cells; AA4-depleted fetal liver cells (FLAA4-); AA4-enriched fetal liver cells (FLAA4+); AA4-depleted bone marrow cells (BMAA4-); AA4-enriched bone marrow cells (BMAA4+). Indicated on the right are positions of molecular weight markers.
20

FIGURES 8A-8C. AA4 expression in retrovirus infected cells. Figure 8A. Flow cytometry analysis of NIH 3T3 and EML Cl cells using PE-conjugated AA4.1 antibody before and after infection with REBNA/AA4. Figure 8B. Immunoprecipitation of biotin-labeled surface proteins using rat IgG2b isotype control antibody (lanes 1, 3, 5) or
25 AA4.1 mAb (lanes 2, 4, 6). Cellular extracts were prepared from D2N cells (lanes 1 and 2), REBNA/AA4 infected NIH 3T3 fibroblasts (lanes 3 and 4), and REBNA/AA4 infected EML Cl cells (lanes 5 and 6). Figure 8C. Immunoprecipitation of cellular extracts using AA4.1. REBNA/AA4 NIH 3T3 cells (lanes 2 -- 4) and REBNA/AA4 EML Cl cells (lanes 6 -- 8) were labeled with 35S-methionine and chased with nonradioactive
30 media for 10 min. (lanes 3 and 7) or 20 min. (lanes 4 and 8). REBNA/GFP NIH 3T3 (lane 1) and REBNA/GFP EML Cl cells (lane 5) are shown as controls. Indicated on the right are positions of molecular weight markers.

FIGURE 9. Nucleotide sequence (SEQ.ID.No.: 72) and the deduced amino acid sequence of AA4 (SEQ.ID.No.: 73). Amino acid sequences is numbered 1 with initiator methionine. Underlined is the putative signal peptide in AA4. The transmembrane domain is double underlined. Light grey line indicates C-type lectin carbohydrate recognition domain. Dark grey lines show EGF-like repeats, while broken grey lines indicate calcium-binding EGF motifs.

FIGURES 10A-10B. Comparison of the primary structures of AA4 and human ClqR.

Figure 10A. Alignment of amino acid sequences of AA4 (SEQ.ID.No.: 73) and ClqR.

Figure 10B. Comparison of the domain structures of AA4 and ClqR. The proteins contain N-terminal signal peptides (SP), a C-type lectin recognition domain, six EGF-like domains, including three calcium-binding (cb) EGF-like domains, and a transmembrane domain (TM).

FIGURE 11. In normal tissues and transformed cell lines, a 7kb RNA species hybridizes with the cloned cDNA (Figures 11A and 11B). In addition to the 7kb species, poly(A)-RNA from D2N cells contains a minor band corresponding to a 3.2 kb mRNA (Figure 11A, lane 8).

FIGURES 12A-12B. Northern blot analysis of AA4 expression in transformed cell lines **Figure 12A** and normal mouse tissues **Figure 12B.** Indicated on the left are positions of 9.44 - 0.24 kb RNA molecular weight markers. Hybridization with D2N poly(A)+ RNA is shown after a 2 hr and overnight exposures. Hybridizations with GAPDH are shown as controls for equal loading.

25

FIGURES 13A-13C. RT-PCR of cDNAs prepared from murine fetal liver, (A) bone marrow-derived hematopoietic cells (B), or differentiating ES cells (C). **Figure 13A.** Lanes 1 and 2, AA4- cells; lane 3, AA4+ cells; lane 4, AA4+Sca-l+c-Kit+LinIO cells; lane 5, AA4+Sca-l-c-Kit+LinIO cells. **Figure 13B.** Lane 1, RhoIOSca-l+Thy-11OLincells; lane 2, RhohiSca-l+Thy-11OLin- cells; lane 3, Sca-l+Thy-l-Lin- cells; lane 4, Lin+ cells. **Figure 13C.** ESO, undifferentiated ES cells; BL1 and 2, blast cell colonies; ENT1 and 2, differentiated endothelial cells; HMT1, 2 and 3, differentiated hematopoietic cells. The cDNAs probes used for hybridization are indicated on the right.

30

FIGURE 14. AFT024 maintains *in vivo* repopulating stem cells. The ability of 3 different stromal cell lines to support highly purified fetal liver stem cells was studied. Freshly purified day 14 fetal liver cells were transplanted directly (10^3 AA4.1⁺, lin^{-lo}, Sca-1⁺, c-kit⁺ cells plus 10^6 Ly5.1 competitor BM per mouse; n=6) into Ly5.1 congenic mice (Control). Additional stem cells from the same purification were also used to initiate Dexter-LTC over irradiated AFT024, 2012, and 2018 stromal cell monolayers (10^4 cells/10 cm dish). After 4 weeks of culture, 10% aliquots of each culture (10^3 stem cell equivalents) were transplanted into Ly5.1 recipients (n=8) together with 10^6 competitor Ly5.1 BM cells. Peripheral blood cells from mice were analyzed for the presence of Ly5.2⁺ cells at 5, 12, 24, and 56 weeks after transplant. Error bars represent SEM.

FIGURE 15. Long-term culture on AFT024 maintains greater levels of repopulating stem cell activity than short-term cytokine- or short-term AFT024-supported cultures. The levels of stem cell activity maintained in short-term cytokine-supported and short-term AFT024-supported cultures were compared to those maintained in long-term AFT024 coculture. Purified fetal liver cells were cultured for 5 days in suspension with cytokines or on an AFT024 monolayer (3000/well-12-well tray). Additional cells from the same purification were seeded onto AFT024 monolayers (3000/6 cm dish) and maintained in Dexter-LTC for 5 weeks. At completion of both culture periods the cells were harvested, mixed with Ly5.1 BM and used to transplant mice. Each mouse received 20% of each culture (600 stem cell equivalents) and 4×10^5 competitor BM cells (4 mice/culture). Peripheral blood cells from mice were analyzed for the presence of Ly5.2⁺ cells at 15 weeks after transplant. FL 1.0 ± 0.57 ; FL/SL 0.75 ± 0.25 ; FL/IL-6 1.8 ± 0.14 ; SL/IL-6 3.2 ± 0.46 ; FL/IL-6/SL 1.7 ± 0.21 ; AFT024 5 days 2.8 ± 0.11 ; AFT024 5 weeks 32.2 ± 7.4 * ($p < 0.004$) Students t-test. See Methods for culture conditions and cytokine concentrations. FL, flk2/flt3-ligand; SL, steel factor. Error bars represent SEM.

FIGURE 16. Time course of cobblestone area formation on AFT024. The formation of stromal dependent CA derived from purified fetal liver stem cells was studied in AFT024 cocultures. Characteristic clusters of at least 50 cells were scored as CA over 28 days of culture. Results are expressed as the mean number of CA/1000 input stem cells from 3

separate fetal liver purifications (300-600 cells/well in 12-well trays). Error bars represent the SEM. The frequency of CA after 28 days is approximately 1 for every 20 input stem cells.

- 5 **FIGURE 17.** High-proliferative potential multilineage clonogenic progenitors are selectively expanded on AFT024. The clonogenic progenitor content of stem cells maintained in AFT024 supported Dexter-LTC was determined. Enriched fetal liver stem cells were seeded onto AFT024 monolayers, at various time points, an individual well was harvested and the cells placed into semi-solid clonogenic progenitor assay (CFU-C).
- 10 Colonies were scored at 8-14 days. Colonies were designated as HPP upon reaching a size ≥ 1 mm after 8 days. CFU numbers at days 0, 4, and 28 are averaged from 3-5 individual stem cell purifications. Error bars represent the SEM for these experiments. Other time points are individual determinations. CFU are normalized to 1000 input stem cells in the stromal cocultures for comparison to day 0 progenitors. *CFU-Mix ($p=.01$) and *CFU-HPP-Mix ($p=.001$) are significantly expanded at day 28 compared to day 0, (Student's T-test).

- FIGURES 18A-18B.** Cobblestone area-initiating cells are expanded on AFT024. **Figure 18A.** A quantitative estimate of the number of 28 day cultured stem cell equivalents required to form a CA after replating on secondary AFT024 monolayers was determined. Four different AFT024 cocultures from separate fetal liver purifications were studied (A, B, C, and D). In limiting-dilution assay, the frequency of stem cell equivalents required to form a CA in another 7 days was 1 in 4 (3.56 ± 0.64 , $r^2=0.96$).
- 20 **Figure 18 B.** CA maintenance in one of the 4 above cultures was followed for an additional 4 weeks. The frequency of stem cell equivalents maintaining CA was determined. At 37% negative wells the frequencies were: 2 weeks 1 in 3, 3 weeks 1 in 10, 4 weeks 1 in 19, and at 5 weeks 1 in 29.

- FIGURES 19A-19B.** Dlk expression analysis in stromal cell lines. **Figure 19A.** (Top) A 1.6 kb dlk transcript is visualized in the parental AFT024 and 2012 cell lines and their subclones, but not in 2018, CFC034 and BFC012. (Bottom) The same filter hybridized with a b-actin probe. **Figure 19B.** RT-PCR analysis of 14 fetal liver-derived stromal cell lines and other cell lines.
- 30

FIGURE 20. Cobblestone area formation by hematopoietic stem cells in the presence of soluble dlk protein. Data are from 4 experiments; 2 each with adult BM cells (Sca-1⁺, c-kit⁺, lin^{lo/-}) and day 14 fetal liver cells (AA4.1⁺, Sca-1⁺, c-kit⁺, lin^{lo/-}). Results are expressed as the ratio/fold-increase in CSA number for fourteen data points each (bars represent, none vs control; dlk vs none; dlk vs control) for the 4 different experiments. Error bars represent SEM. *P= 0.01 comparing dlk vs none to none vs control, **P= 0.001 comparing dlk vs control to none vs control (Student's t-test).

FIGURE 21. Membrane bound dlk expression in transfected BFC012 cells. Full length dlk cDNA was transfected into BFC012 cells. (*Left*) A flow diagram of dlk expression in transfected BFC012 populations (BFC-dlk) and cells transfected with the selection plasmid alone (BFC-Zeo). (*Right*) Expression of dlk in a cloned line (BFC-dlk-5) derived from the expressing population and a control clone (BFC-Zeo-1).

FIGURES 22A. CSA formation by hematopoietic stem cells in the presence of membrane-bound dlk. **Figure 22A.** Bars labeled BFC are from 5 groups (nontransfected BFC012 cells, two control pZeo transfected BFC012 populations, and two clones derived from the pZeo transfected populations). Bars labeled BFC-dlk are from three groups shown to express transfected dlk; one dlk-transfected BFC012 population (BFC-dlk) and two individual transfected clones (BFC-dlk-1 and BFC-dlk-5). Error bars represent SEM. ** P< 0.001 days 3, 4, and 5; *P< 0.01 days 6 and 7 (Student's t-test). **Figure 22B.** A clone derived from the dlk transfected populations of BFC012 cells (BFC-dlk-5) and a clone derived from pZeo transfected populations (BFC-Zeo-1) were used for CSA assay with purified fetal liver stem cells. CSA/1000 input stem cells are expressed as the mean of three individual experiments, error bars represent the SEM. **P< 0.001 at days 4, 6, and 8 (Student's t-test).

FIGURES 23A-23B. HPP multilineage clonogenic progenitors and *in vivo* repopulating stem cells are maintained in short-term dlk-expressing cocultures. **Figure 23A.** Fetal liver stem cells were purified as described and assayed for their progenitor content immediately after purification and after culture on BFC-dlk-5, and BFC-Zeo-1. At day 4 the cultures were used for clonogenic progenitor (3 experiments) and transplantation assay (2 experiments). Bars represent data from 3 experiments with day 0 cells (Fresh) and day 4

cocultured cells (BFC-dlk-5 and BFC-Zeo-1), error bars represent SEM. *P= 0.01 for total CFU-C from fresh stem cells compared to total CFU-C from BFC-dlk-5 cocultures at day 4; ** P= 0.001 for total CFU-C from BFC-dlk-5 compared to total CFU-C from BFC-Zeo-1 (Student's t-test). **Figure 23 B.** Analysis of *in vivo* repopulating ability of purified fetal liver stem cells cocultured for 4 days on BFC-dlk-5, BFC-Zeo-1, and AFT024 monolayers. Results are from nine individual mice in two experiments (4-5 mice in each experiment) at 10 weeks after transplantation. P= 0.05 for BFC-dlk-5 vs BFC-Zeo-1 (Student's t-test).

FIGURE 24. Genes and predicted proteins isolated from primitive stem cells by the techniques of the present invention. (Sequence Identification Numbers are indicated on the figure).

DETAILED DESCRIPTION

The present invention provides an isolated nucleic acid derived from an isolated hematopoietic stem cell, the isolated nucleic acid comprising the following characteristics: (1) specifically expressed in the hematopoietic stem cell; and (2) encoding a hematopoietic stem cell – specific protein.

An embodiment of this invention further comprises the following characteristic:
capable of hybridizing under standard conditions with a sequence selected from the group consisting of SEQ.ID.No.: 1, SEQ.ID.No.: 2, SEQ.ID.No.: 3, SEQ.ID.No.: 4, SEQ.ID.No.: 5, SEQ.ID.No.: 6, SEQ.ID.No.: 7, SEQ.ID.No.: 8, SEQ.ID.No.: 9, SEQ.ID.No.: 10, SEQ.ID.No.: 11, SEQ.ID.No.: 12, SEQ.ID.No.: 13, SEQ.ID.No.: 14, SEQ.ID.No.: 15, SEQ.ID.No.: 16, SEQ.ID.No.: 17, SEQ.ID.No.: 18, SEQ.ID.No.: 19, SEQ.ID.No.: 20, SEQ.ID.No.: 21, SEQ.ID.No.: 22, SEQ.ID.No.: 23, SEQ.ID.No.: 24, SEQ.ID.No.: 25, SEQ.ID.No.: 26, SEQ.ID.No.: 27, SEQ.ID.No.: 28, SEQ.ID.No.: 29, SEQ.ID.No.: 30, SEQ.ID.No.: 31, SEQ.ID.No.: 32, SEQ.ID.No.: 33, SEQ.ID.No.: 34, SEQ.ID.No.: 35, SEQ.ID.No.: 36, SEQ.ID.No.: 37, SEQ.ID.No.: 38, SEQ.ID.No.: 39, SEQ.ID.No.: 40, SEQ.ID.No.: 41, SEQ.ID.No.: 42, and SEQ.ID.No.: 43, SEQ.ID.No.: 45, SEQ.ID.No.: 47, SEQ.ID.No.: 49, SEQ.ID.No.: 51, SEQ.ID.No.: 53, SEQ.ID.No.: 55, SEQ.ID.No.: 57, SEQ.ID.No.: 59, SEQ.ID.No.: 61, SEQ.ID.No.: 63, SEQ.ID.No.: 65, SEQ.ID.No.: 67, SEQ.ID.No.: 72 or a portion thereof. A portion thereof, in a preferred embodiment of this invention is the 5' end region or the 3' end region of the nucleic acid.

In another preferred embodiment, a portion thereof is at least a 8-18 nucleotide portion of the coding region. In yet another preferred embodiment, a portion thereof is at least a 8-18 nucleotide portion of a non-coding regulatory region or a binding region such as a stem-cell specific promoter or enhancer region. According to still another embodiment of the provided isolated nucleic acid further comprises the characteristic of encoding a protein capable of modulating hematopoietic stem cell activity. According to this invention modulating hematopoietic stem cell activity includes up-regulating, down-regulating or otherwise changing the activity of the hematopoietic stem cell. Such activity is contemplated as inducing differentiation or inhibiting differentiation of the cell.

However, directing differentiation toward one or another daughter cell type is also within the scope of a preferred embodiment of this invention. Other preferred embodiments include but are not limited to modulation of transcription, translation, gene splicing, transport, proteolytic processing, replication, expression of cell surface markers and transplantation. According to still another embodiment of the present invention, the activity is selected from the group consisting of hematopoietic stem cell differentiation and hematopoietic stem cell replication. According to yet another embodiment of this invention, the protein is selected from the group consisting of a growth factor, a transcription factor, a splicing factor, a capping factor, a transport protein, a translation factor, and a replication factor. In one preferred embodiment of this invention, the provided nucleic acid comprises the nucleotide sequence of SEQ.ID.No.: 72, an analog thereof, or a portion thereof. According to another preferred embodiment of this invention, the hematopoietic stem cell is a primitive hematopoietic stem cell. In one embodiment of this invention, the primitive hematopoietic stem cell is selected from the group consisting of an umbilical cord cell, a bone marrow cell and a fetal liver cell. In a preferred embodiment of this invention, the primitive hematopoietic stem cell is selected from the group consisting of a AFT024 cell, a 2012 cell and a 2018 cell.

The present invention further provides a composition comprising the provided nucleic acid, wherein the nucleic acid comprises one selected from the group consisting of SEQ.ID.No.: 1, SEQ.ID.No.: 2, SEQ.ID.No.: 3, SEQ.ID.No.: 4, SEQ.ID.No.: 5, SEQ.ID.No.: 6, SEQ.ID.No.: 7, SEQ.ID.No.: 8, SEQ.ID.No.: 9, SEQ.ID.No.: 10, SEQ.ID.No.: 11, SEQ.ID.No.: 12, SEQ.ID.No.: 13, SEQ.ID.No.: 14, SEQ.ID.No.: 15, SEQ.ID.No.: 16, SEQ.ID.No.: 17, SEQ.ID.No.: 18, SEQ.ID.No.: 19, SEQ.ID.No.: 20, SEQ.ID.No.: 21, SEQ.ID.No.: 22, SEQ.ID.No.: 23, SEQ.ID.No.: 24, SEQ.ID.No.: 25,

SEQ.ID.No.: 26, SEQ.ID.No.: 27, SEQ.ID.No.: 28, SEQ.ID.No.: 29, SEQ.ID.No.: 30,
SEQ.ID.No.: 31, SEQ.ID.No.: 32, SEQ.ID.No.: 33, SEQ.ID.No.: 34, SEQ.ID.No.: 35,
SEQ.ID.No.: 36, SEQ.ID.No.: 37, SEQ.ID.No.: 38, SEQ.ID.No.: 39, SEQ.ID.No.: 40,
SEQ.ID.No.: 41, SEQ.ID.No.: 42, SEQ.ID.No.: 43, SEQ.ID.No.: 45, SEQ.ID.No.: 47,
5 SEQ.ID.No.: 49, SEQ.ID.No.: 51, SEQ.ID.No.: 53, SEQ.ID.No.: 55, SEQ.ID.No.: 57,
SEQ.ID.No.: 59, SEQ.ID.No.: 61, SEQ.ID.No.: 63, SEQ.ID.No.: 65, SEQ.ID.No.: 67,
SEQ.ID.No.: 72 or a portion thereof. According to one embodiment of this invention, the
the nucleic acid is selected from the group consisting of DNA, RNA and cDNA. Another
embodiment of this invention is a vector comprising the provided nucleic acid.
10 According to yet another embodiment, the vector comprises viral or plasmid DNA. A
further embodiment of this invention is an expression vector comprising the provided
nucleic acid and a regulatory element. A still further embodiment of this invention is a
host vector system which comprises the expression vector in a suitable host. In a
preferred embodiment of this invention, the suitable host is selected from the group
15 consisting of a bacterial cell, a eukaryotic cell, a mammalian cell and an insect cell.

The present invention additionally provides an isolated hematopoietic stem cell
specific protein or a portion thereof encoded by the provided nucleic acid. According to
one embodiment of this invention is the provided protein further comprising the
characteristic of being capable of modulating hematopoietic stem cell activity. According
20 to this invention modulating hematopoietic stem cell activity includes up-regulating,
down-regulating or otherwise changing the activity of the hematopoietic stem cell. Such
activity is contemplated as inducing differentiation or inhibiting differentiation of the cell.
However, directing differentiation toward one or another daughter cell type is also within
the scope of a preferred embodiment of this invention. Other preferred embodiments
25 include but are not limited to modulation of transcription, translation, gene splicing,
transport, proteolytic processing, replication, expression of cell surface markers and
transplantation. According to still another embodiment of the present invention, the
activity is selected from the group consisting of hematopoietic stem cell differentiation
and hematopoietic stem cell replication. According to yet another embodiment of this
30 invention, the protein is selected from the group consisting of a growth factor, a
transcription factor, a splicing factor, a capping factor, a transport protein, a translation
factor, and a replication factor. According to one embodiment the activity is selected
from the group consisting of hematopoietic stem cell differentiation and hematopoietic

stem cell replication. According to another embodiment, the protein is selected from the group consisting of a growth factor, a transcription factor, a splicing factor, a capping factor, a transport protein, a translation factor, and a replication factor. According to still another embodiment, the protein has substantially the same amino acid sequence as one
5 selected from the group consisting of SEQ.ID.No.: 42, SEQ.ID.No.: 44, SEQ.ID.No.: 46, SEQ.ID.No.: 48, SEQ.ID.No.: 50, SEQ.ID.No.: 52, SEQ.ID.No.: 54, SEQ.ID.No.: 56, SEQ.ID.No.: 58, SEQ.ID.No.: 60, SEQ.ID.No.: 62, SEQ.ID.No.: 64, SEQ.ID.No.: 66, SEQ.ID.No.: 68, SEQ.ID.No.: 70, SEQ.ID.No.: 71, and SEQ.ID.No.: 73.

The present invention further still provides a nucleic acid probe capable of
10 specifically hybridizing with the provided nucleic acid under standard hybridization conditions.

Also, the present invention provides an antibody capable of specifically binding to the provided protein without substantially cross-reacting with a non-stem cell specific protein or homologs thereof under conditions permissive to antibody binding.
15 Additionally, the present invention provides a cell capable of producing the provided antibody.

In addition, the present invention provides a method for identifying the presence of a primitive hemopoietic stem cell in a sample comprising nucleic acids specifically expressed in hematopoietic stem cells comprising (a) contacting the sample with the
20 provided antibody under conditions permissive to the formation of an antibody complex; (b) detecting the presence of the complex formed in step (a), the presence of a complex formed indicating the presence of a primitive hemopoietic stem cell in the sample. According to one embodiment of this invention, the antibody is labeled with a detectable marker. In a preferred embodiment, the detectable marker is selected from the group
25 consisting of a radioactive isotope, enzyme, magnetic bead, dye, fluorescent marker and biotin.

Further still, the present invention provides a method for generating a stem cell/progenitor cell from a primitive hematopoietic cell in a sample comprising contacting the sample with the provided protein. Another embodiment of this invention, provides a
30 method for generating a stem cell/progenitor cell from a primitive hematopoietic cell in a sample comprising contacting the sample with the provided nucleic acid. According to an preferred embodiment, the nucleic acid is in an expression vector. According to another preferred embodiment the nucleic acid is introduced into the cell under conditions

permissive to the expression of the nucleic acid.

The present invention further provides a method for identifying the presence in a sample of a compound that modulates hematopoietic stem cell activity comprising: (a) contacting the hematopoietic stem cell with the sample; (b) determining the hematopoietic stem cell activity; and (c) comparing the hematopoietic stem cell activity determined in step (b) with the activity determined in the absence of the compound an increase or decrease in hematopoietic stem cell activity indicating the presence in the sample of a compound that modulates hematopoietic stem cell activity. According to one embodiment, the activity is selected from the group consisting of gene expression, replication, differentiation, transplantation, and self regeneration. The present invention also still further provides a compound identified by the method of this invention, previously unknown.

The present invention even further still provides a method for identifying primitive hematopoietic stem cell-specific nucleic acids , comprising: (a) creating a primitive hematopoietic stem cell cDNA library and a non-primitive stem cell immune cell cDNA library; and (b) subtracting the two libraries, thereby identifying primitive stem cell specific nucleic acids. According to one embodiment is (i) contacting the nucleic acids of the stem cell and non-stem cell libraries with each other under conditions permissive to hybridization, thereby forming hybrid complexes; (ii) separating the hybrid complexes formed in step (b) from the nucleic acids which did not form complexes; and (iii) isolating the nucleic acids which did not form complexes, thereby identifying hematopoietic stem cell specific nucleic acids. In still another embodiment, step (ii) further comprising amplification of the nucleic acids. Yet another embodiment is step (iii) further comprising amplication of the nucleic acids which did not form complexes. Even still another embodiment is further comprising displaying the amplified DNA on a chromatography gel. A further embodiment is step (b) comprising differential display of the two libraries, thereby identifying primitive stem cell specific nucleic acids. Also yet another embodiment is step (b) comprising representation difference analysis of the two libraries, thereby identifying primitive stem cell specific nucleic acids. Yet even another embodiment is further comprising cloning the stem cell specific nucleic acid. According to a preferred embodiment, the stem cell is selected from the group consisting of AF024, 2012, and 2018. The present invention further provides a nucleic acid identified by the provided method.

The present invention additionally provides a composition comprising the provided compound and a carrier. An embodiment of this invention is a pharmaceutical composition comprising the provided compound and a pharmaceutically acceptable carrier.

5 Also the present invention additionally provides a primitive hematopoietic stem cell specifically expressing one selected from the group consisting of: SEQ.ID.No.: 1, SEQ.ID.No.: 2, SEQ.ID.No.: 3, SEQ.ID.No.: 4, SEQ.ID.No.: 5, SEQ.ID.No.: 6, SEQ.ID.No.: 7, SEQ.ID.No.: 8, SEQ.ID.No.: 9, SEQ.ID.No.: 10, SEQ.ID.No.: 11, SEQ.ID.No.: 12, SEQ.ID.No.: 13, SEQ.ID.No.: 14, SEQ.ID.No.: 15, SEQ.ID.No.: 16, 10 SEQ.ID.No.: 17, SEQ.ID.No.: 18, SEQ.ID.No.: 19, SEQ.ID.No.: 20, SEQ.ID.No.: 21, SEQ.ID.No.: 22, SEQ.ID.No.: 23, SEQ.ID.No.: 24, SEQ.ID.No.: 25, SEQ.ID.No.: 26, SEQ.ID.No.: 27, SEQ.ID.No.: 28, SEQ.ID.No.: 29, SEQ.ID.No.: 30, SEQ.ID.No.: 31, SEQ.ID.No.: 32, SEQ.ID.No.: 33, SEQ.ID.No.: 34, SEQ.ID.No.: 35, SEQ.ID.No.: 36, SEQ.ID.No.: 37, SEQ.ID.No.: 38, SEQ.ID.No.: 39, SEQ.ID.No.: 40, SEQ.ID.No.: 41, 15 SEQ.ID.No.: 42, and SEQ.ID.No.: 43, SEQ.ID.No.: 45, SEQ.ID.No.: 47, SEQ.ID.No.: 49, SEQ.ID.No.: 51, SEQ.ID.No.: 53, SEQ.ID.No.: 55, SEQ.ID.No.: 57, SEQ.ID.No.: 59, SEQ.ID.No.: 61, SEQ.ID.No.: 63, SEQ.ID.No.: 65, SEQ.ID.No.: 67, SEQ.ID.No.: 72 or a portion thereof. An embodiment of this invention is a primitive hematopoietic stem cell specifically expressing a nucleic acid identified by the provided method.

20 Yet additionally, the present invention provides a method for treating a condition in a subject comprising administering to the subject a therapeutically effective amount of the provided pharmaceutical composition. According to an embodiment of this invention, the condition is an immune system condition. In a further embodiment of this invention, the condition is leukemia.

25 The present invention provides a method of introducing an exogenous nucleic acid into a hematopoietic stem cell comprising contacting the stem cell with the provided composition.

 Finally, the present invention provides a method of ex vivo expansion of hematopoietic stem cells comprising contacting the cell with the provided composition. 30 According to an embodiment of this invention, the ex vivo expanded hematopoietic stem cells are available for therapeutic use. The expanded cells are available to receive exogenous genes, including by retroviral or other vectors which require a round of

replication. Alternatively, the cells are available for transplantation either autologously or heterologously.

As used herein, the term, a sequence is conserved if there is substantial homology of sequence between multiple gene species.

5 As used herein, the terms, "hybridization" and "binding" in the context of probes, primers and denatured DNA are used interchangeably. Probes which are hybridized or bound to denatured DNA are aggregated to complementary sequences in the polynucleotide. Whether or not a particular probe remains aggregated with the polynucleotide depends on the degree of complementarity, the length of the probe, and
10 the stringency of the binding conditions. The higher the stringency, the higher must the degree of complementarity, and/or the longer the probe.

As used herein, the term, probe, refers to an oligonucleotide designed to be sufficiently complementary to a sequence in a denatured nucleic acid to be probed, in relation to its length, to be bound under selected stringency conditions. Primers may vary
15 in length. Preferably such primers should be sufficiently long to hybridize to the modified RNAs in a specific and stable manner. A semi-random primer as the term is used herein, encompasses a class of primers wherein either a discrete portion of the primer is random, while another discrete portion is conserved as well as primers which have nucleotide preferences at particular positions within a sequence. For example, the
20 discrete portion-type primer may have a predetermined adaptor sequence at its 5' end and a random sequence at its 3' end. Alternatively, several preferred primers have nucleotide preferences at specific positions within the primers while other positions are random. A degenerate primer as the term is used herein, encompasses a cocktail or mixture of primers wherein one or more of the possible triplet nucleotide sequences encoding an
25 amino acid is incorporated into the primer sequence. For example, Serine may be encoded by six separate triple sequences (AGU, AGC, UCU, UCC, UCA, and UCG). Thus, a "degenerate" primer may reflect the degeneracy of the nucleotide triplet code. Alternatively, a randomized primer, as the term is used herein, encompasses a primer wherein, the nucleotide at one or more positions may be randomized in order to yield a
30 triplet sequence encoding an alternative or a random amino acid at the position.

An end region as the term is used herein, consists of the end nucleotide and a portion of the region including as much as that half of the entire sequence. For example, the "3' end region" or "3' region" of a primer may include the 3' half of the primer.

A preferred method of hybridization is blot hybridization. See Sambrook et al. 1989 *Molecular Cloning: A Laboratory Manual* 2nd Ed. for additional details regarding blot hybridization. Using this method, separated amplification products are transferred onto a solid matrix, such as a filter. The probe, which is detectable, either directly or
5 indirectly, is hybridized to the solid matrix under appropriate conditions. The matrix is washed to remove excess probe. Thereafter the probe which specifically hybridized to the solid matrix can be detected.

The probe can be DNA or RNA and can be made detectable by any of the many labeling techniques readily available and known to the skilled artisan. Such methods
10 include, but are not limited to, radio-labeling, digoxigenin-labeling, and biotin-labeling. A well-known method of labeling DNA is ^{32}P using DNA polymerase, Klenow enzyme or polynucleotide kinase. In addition, there are known non-radioactive techniques for signal amplification including methods for attaching chemical moieties to pyrimidine and purine rings (Dale, R.N.K. et al, 1973 *Proc. Natl. Acad. Sci. USA* 70:2238-42), methods
15 which allow detection by chemiluminescence (Barton, S.K. et al, 1992 *J. Am. Chem. Soc.* 114:8736-40) and methods utilizing biotinylated nucleic acid probes (Johnson, T.K. et al, 1983 *Anal. Biochem.* 133:125-131; Erickson, P.F. et al, 1982 *J. Immunol. Methods* 51:241-49; Matthaei, F.S. et al, 1986 *Anal. Biochem.* 157:123-28) and methods which allow detection by fluorescence using commercially available products. Non-radioactive
20 labeling kits are also commercially available.

A basic description of nucleic acid amplification is described in Mullis, U.S. Patent No. 4,683,202, which is incorporated herein by reference. The amplification reaction uses a template nucleic acid contained in a sample, two primer sequences and inducing agents. The extension product of one primer when hybridized to the second
25 primer becomes a template for the production of a complementary extension product and vice versa, and the process is repeated as often as is necessary to produce a detectable amount of the sequence.

The inducing agent may be any compound or system which will function to accomplish the synthesis of primer extension products, including enzymes. Suitable
30 enzymes for this purpose include, for example, *E.coli* DNA polymerase I, thermostable *Taq* DNA polymerase, Klenow fragment of *E.coli* DNA polymerase I, T4 DNA polymerase, other available DNA polymerases, reverse transcriptase and other enzymes which will facilitate combination of the nucleotides in the proper manner to form

amplification products. The oligonucleotide primers can be synthesized by automated instruments sold by a variety of manufacturers or can be commercially prepared based upon the nucleic acid sequence of this invention.

5 The degree of hybridization depends on the degree of complementarity, the length of the nucleic acid molecules being hybridized, and the stringency of the conditions in a reaction mixture. Stringency conditions are affected by a variety of factors including, but not limited to temperature, salt concentration, concentration of the nucleic acids, length of the nucleic acids, sequence of the nucleic acids and viscosity of the reaction mixture. More stringent conditions require greater complementarity between the nucleic acids in
10 order to achieve effective hybridization.

"Hybridization" and "binding" in the context of probes, primers and denatured DNA are used interchangeably. Probes which are hybridized or bound to denatured DNA are aggregated to complementary sequences in the polynucleotide. Whether or not a particular probe remains aggregated with the polynucleotide depends on the degree of
15 complementarity, the length of the probe, and the stringency of the binding conditions. The higher the stringency, the higher must the degree of complementarity, and/or the longer the probe.

"Probe" refers to an oligonucleotide designed to be sufficiently complementary to a sequence in a denatured nucleic acid to be probed, in relation to its length, to be bound
20 under selected stringency conditions.

Primers may vary in length. Preferably such primers should be sufficiently long to hybridize to the modified RNAs in a specific and stable manner.

A semi-random primer as the term is used herein, encompasses a class of primers wherein either a discrete portion of the primer is random, while another discrete portion is
25 conserved as well as primers which have nucleotide preferences at particular positions within a sequence. For example, the discrete portion-type primer may have a predetermined adaptor sequence at its 5' end and a random sequence at its 3' end. Alternatively, several preferred primers have nucleotide preferences at specific positions within the primers while other positions are random.

30 A degenerate primer as the term is used herein, encompasses a cocktail or mixture of primers wherein one or more of the possible triplet nucleotide sequences encoding an amino acid is incorporated into the primer sequence. For example, Serine may be encoded by six separate triple sequences (AGU, AGC, UCU, UCC, UCA, and UCG).

Thus, a "degenerate" primer may reflect the degeneracy of the nucleotide triplet code. Alternatively, a randomized primer, as the term is used herein, encompasses a primer wherein, the nucleotide at one or more positions may be randomized in order to yield a triplet sequence encoding an alternative or a random amino acid at the position.

5 Solid matrices are available to the skilled artisan. Solid phases useful to serve as a matrix for the present invention include but are not limited to polystyrene, polyethylene, polypropylene, polycarbonate, or any solid plastic material in the shape of test tubes, beads, microparticles, dip-sticks, plates or the like. Additionally matrices include, but are not limited to membranes, 96-well microtiter plates, test tubes and Eppendorf tubes.
10 Solid phases also include glass beads, glass test tubes and any other appropriate shape made of glass. A functionalized solid phase such as plastic or glass which has been modified so that the surface carries carboxyl, amino, hydrazide, or aldehyde groups can also be used. In general such matrices comprise any surface wherein a ligand-binding agent can be attached or a surface which itself provides a ligand attachment site.

15 As used herein, "pharmaceutically acceptable" refers to molecular entities and compositions that are physiologically tolerable and do not typically produce an allergic or similar untoward reaction, such as gastric upset, dizziness and the like, when administered to a human. A pharmaceutically acceptable carrier encompasses any of the standard pharmaceutically accepted carriers, such as phosphate buffered saline
20 solution, water emulsions such as an oil/water emulsion or a triglyceride emulsion, various types of wetting agents, tablets, coated tablets and capsules. Typically such carriers contain excipients such as starch, milk, sugar, certain types of clay, gelatin, stearic acid, talc, vegetable fats or oils, gums, glycols, or other known excipients. Such carriers may also include flavor and color additives or other ingredients. The invention
25 also provides for pharmaceutical compositions together with suitable diluents, preservatives, solubilizers, emulsifiers and adjuvants. Other embodiments of the compositions of the invention incorporate particulate forms, protective coatings, protease inhibitors or permeation enhancers for various routes of administration, including but not limited to intravenous, intramuscular, parenteral, pulmonary, nasal
30 and oral.

 As used herein, an "effective amount" is the amount required to achieve a clinically significant effect. For example a significant reduction of infection, or

reduction of cell growth or reduction of tumor size is a reduction of preferably of at least 30 percent, more preferably of at least 50 percent, most preferably of at least 90 percent. Accordingly, the effective amount will vary with the subject being treated, as well as the condition to be treated. For the purposes of this invention, the methods of administration are to include, but are not limited to administration cutaneously, subcutaneously, intravenously, parenterally, orally, topically, or by aerosol.

The present invention further contemplates therapeutic compositions useful in practicing the therapeutic methods of this invention. A subject therapeutic composition includes, in admixture, a pharmaceutically acceptable excipient (carrier) and one or more of a polypeptide analog or fragment of the provided peptide or peptide composition, a peptidomimetic composition thereof as described herein as an active ingredient. A cocktail of the provided pharmaceutical composition in various combinations is also contemplated.

The preparation of therapeutic compositions which contain polypeptides, analogs or active fragments as active ingredients is well understood in the art. Typically, such compositions are prepared as injectables, either as liquid solutions or suspensions, however, solid forms suitable for solution in, or suspension in, liquid prior to injection can also be prepared. The preparation can also be emulsified. The active therapeutic ingredient is often mixed with excipients which are pharmaceutically acceptable and compatible with the active ingredient. Suitable excipients are, for example, water, saline, dextrose, glycerol, ethanol, or the like and combinations thereof. In addition, if desired, the composition can contain minor amounts of auxiliary substances such as wetting or emulsifying agents, pH buffering agents which enhance the effectiveness of the active ingredient.

A polypeptide, analog or active fragment can be formulated into the therapeutic composition as neutralized pharmaceutically acceptable salt forms. Pharmaceutically acceptable salts include the acid addition salts (formed with the free amino groups of the polypeptide or antibody molecule) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed from the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium,

calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, 2-ethylamino ethanol, histidine, procaine, and the like.

The therapeutic polypeptide-, analog- or active fragment-containing compositions are conventionally administered intravenously, as by injection of a unit
5 dose, for example. The term "unit dose" when used in reference to a therapeutic composition of the present invention refers to physically discrete units suitable as unitary dosage for humans, each unit containing a predetermined quantity of active material calculated to produce the desired therapeutic effect in association with the required diluent; i.e., carrier, or vehicle.

10 The compositions are administered in a manner compatible with the dosage formulation, and in a therapeutically effective amount. The quantity to be administered depends on the subject to be treated, capacity of the subject's immune system to utilize the active ingredient, and degree of inhibition desired. Precise amounts of active ingredient required to be administered depend on the judgment of the practitioner and
15 are peculiar to each individual. However, suitable dosages may range from about 0.1 to 20, preferably about 0.5 to about 10, and more preferably one to several, milligrams of active ingredient per kilogram body weight of individual per day and depend on the route of administration. Suitable regimes for initial administration and booster shots are also variable, but are typified by an initial administration followed by repeated doses at
20 one or more hour intervals by a subsequent injection or other administration. Alternatively, continuous intravenous infusion sufficient to maintain concentrations of ten nanomolar to ten micromolar in the blood are contemplated.

As used herein, the term "synthetic amino acid" means an amino acid which is chemically synthesized and is not one of the 20 amino acids naturally occurring in nature.

25 As used herein, the term "biosynthetic amino acid" means an amino acid found in nature other than the 20 amino acids commonly described and understood in the art as "natural amino acids."

As used herein, amino acid residues are preferred to be in the "L" isomeric form. However, residues in the "D" isomeric form can be substituted for any L-amino
30 acid residue, as long as the desired functional property is retained by the polypeptide. NH₂ refers to the free amino group present at the amino terminus of a polypeptide. COOH refers to the free carboxy group present at the carboxy terminus of a

polypeptide. Abbreviations for amino acid residues are used in keeping with standard polypeptide nomenclature delineated in *J. Biol. Chem.*, 243:3552-59 (1969).

It should be noted that all amino-acid residue sequences are represented herein by formulae whose left and right orientation is in the conventional direction of amino-terminus to carboxy-terminus. Furthermore, it should be noted that a dash at the beginning or end of an amino acid residue sequence indicates a peptide bond to a further sequence of one or more amino-acid residues.

Amino acids with nonpolar R groups include: Alanine, Valine, Leucine, Isoleucine, Proline, Phenylalanine, Tryptophan and Methionine. Amino acids with uncharged polar R groups include: Glycine, Serine, Threonine, Cysteine, Tyrosine, Asparagine and Glutamine. Amino acids with charged polar R groups (negatively charged at Ph 6.0) include: Aspartic acid and Glutamic acid. Basic amino acids (positively charged at pH 6.0) include: Lysine, Arginine and Histidine (at pH 6.0). Amino acids with phenyl groups include: Phenylalanine, Tryptophan and Tyrosine. Particularly preferred substitutions are: Lys for Arg and vice versa such that a positive charge may be maintained; Glu for Asp and vice versa such that a negative charge may be maintained; Ser for Thr such that a free -OH can be maintained; and Gln for Asn such that a free NH₂ can be maintained. Amino acids can be in the "D" or "L" configuration. Use of peptidomimetics may involve the incorporation of a non-amino acid residue with non-amide linkages at a given position.

Amino acid substitutions may also be introduced to substitute an amino acid with a particularly preferable property. For example, a Cys may be introduced a potential site for disulfide bridges with another Cys. A His may be introduced as a particularly "catalytic" site (i.e., His can act as an acid or base and is the most common amino acid in biochemical catalysis). Pro may be introduced because of its particularly planar structure, which induces -turns in the protein's structure.

As used herein, "pM" means picomolar, "nM" means nanmolar, "uM, means micromolar, "mM" means millimolar, "ul" or "μl" mean microliter, "ml" means milliliter, "l" means liter.

The following examples are presented in order to more fully illustrate the preferred embodiments of the invention. They should in no way be construed, however,

as limiting the broad scope of the invention. While the invention is described and illustrated herein by references to various specific material, procedures and examples, it is understood that the invention is not restricted to the particular material combinations of material, and procedures selected for that purpose. Numerous variations of such details 5 can be implied as will be appreciated by those skilled in the art.

EXAMPLES

EXAMPLE 1: Molecules specifically expressed in murine stem/progenitor cells

Protein tyrosine kinases and phosphatases expressed in murine hematopoietic stem cells have been previously identified (Matthews, W., et al. 1991; Matthews, W., et al. 1991; Dosil, M., et al. 1996). These molecules play important roles in hematopoiesis and development (Dosil, M., et al. 1996; Mackarechtschian, K., et al. 1995; Shalaby, F., et al. 1995; Kabrun, N., et al. 1997). The present invention contemplates an even more global approach in order to identify molecules specifically expressed in the murine stem/progenitor cell hierarchy. For most of these studies purified (AA4.1+Lin-/loSca+ckit+) fetal liver cells were used. This population is approximately 1,000-fold enriched for in vivo repopulating activity (LTRA) measured by competitive repopulation (Ly5.1/Ly5.2 congenic system) (Moore, K.A., et al. 1997). All myeloid and lymphoid lineages in primary and in secondary recipients are repopulated by these cells. This degree of enrichment is comparable to the current "state of the art." Other primitive members of the stem/progenitor cell hierarchy share this same cell surface phenotype. These include: (1) LTCIC or cobblestone area forming cells (Ploemacher, R.E., et al. 1989; Ploemacher, R.E., et al. 1991), (2) CFU-blast progenitors, (3) HPP-CFC progenitors (Lowry, P.A., et al. 1995) and (4) stromal-dependent B-lymphoid progenitors (Whitlock, C.A., and Müller-Sieburg, C.E. 1990). The AA4.1+Lin-/loSca-ckit+ subset is depleted of LTRA but contains significant in vitro progenitor activity. In contrast, no stem/progenitor cell activity is found in the AA4.1 subset (Jordan, C.T., et al. 1990). Short-term (5-7 days) cytokine cultures of stem cells were used to generate committed progenitor populations at the expense of LTRA (Traycoff, C.M., et al. 1996; Peters, S.O., et al. 1995; Knobel, K.M., et al. 1994; Yonemura, Y., et al. 1996). In summary, several

cell populations were defined which represent the beginning, the middle and the end points of the hematopoietic hierarchy. This sets the stage for a comparative analysis of gene expression patterns. A goal of the present invention is to complement the physical and functional phenotypes of stem/progenitor cells with profiles of uniquely expressed genes. It was hypothesized that some of these gene products contribute to the unique biological properties of primitive stem/progenitor cells and therefore are regulators of self-renewal, proliferation, commitment and other processes.

There are a number of ways to compare gene expression profiles which are available to one of skill in the art. It is possible to do this by exhaustive sequencing of representative cDNA collections obtained from stem cell and mature cell sources followed by "electronic subtraction". This approach has several drawbacks. Most importantly, the number of sequences which must be obtained is prohibitive. For a homogeneously pure population this number is on the order of 50,000 (based on approximately 10-20,000 expressed genes in an average cell type and a statistical calculation). In practice, even the most purified stem cell population is heterogeneous. Stem cell enrichment values are only meaningful in relation to an unenriched standard and cannot be converted into an absolute stem cell number. It has been documented that cell populations with the same cell-surface phenotypes can differ in biological activity. The unique properties of stem cells also suggest caution when extrapolating from expressed gene numbers in other cells. In short, the extent of sequencing necessary to ensure complete coverage of gene expression in stem cells is not possible to estimate. Normalization procedures (Uchida, N., et al. 1995; Patanjali, S.R., et al. 1991; Soares, M.B., et al. 1994) designed to "equalize" the mRNA abundance classes are not advisable because they obliterate potentially important quantitative expression differences. Additionally, a high-throughput sequencing effort is not applicable to numerous libraries. Comparisons of gene expression in diverse sources of stem cells will provide valuable information. An elegant technique, Serial Analysis of Gene Expression (SAGE) permits the rapid acquisition of thousands of DNA sequence "tags" (Zhang, L., et al. 1997). This technology was not considered herein because the size of each sequence "tag" is very small (10 base pair, bp). Therefore, SAGE is only informative in two extremes; exact nucleotide matches or no matches to sequences in the databases. This limits database comparison to the same species from which the "tags" originate. A key component of the strategy presented herein relies on broad bioinformatic database comparisons. In

addition, even with a specific "tag" one still needs to obtain a full-length cDNA clone for functional studies. Analysis of gene expression can be done in single progenitor cells after these form a colony "start" (Brady, G., et al. 1990; Brady, G., and Iscove, N.N. 1993). Replating of the sibling cells in a "start" colony allows the approximation of the lineage potential present in the starting cell. This technique suffers from several drawbacks. First, it is limited to 3', non-coding ends, thus preventing protein database comparisons. Second, the technique relies on cell growth, thus it is not suitable for analysis of quiescent cells. Moreover, approaches to allow colony-formation by true stem cells with a retention of primitive properties are in their infancy (Ball, T.C., et al. 1995; Trevisan, M., et al. 1996). Third, the technique does not take into account stochastic models of stem cell behavior (Ogawa, M. 1993). To accurately reveal physical and functional properties of stem cells it is wise to analyze populations where stochastic differences would average out. Single-cell derived cDNA populations could however, provide valuable "adjunct" material for more refined gene expression screens. There are many techniques to physically identify differentially expressed genes. For these and other reasons, most notably the technical and economic ease with which physical pre-enrichments of cDNA libraries can be achieved, the studies described herein began with cDNA libraries which are highly enriched in differentially expressed sequences. The integration of individual techniques served to overcome the inherent limitations of each technique. Three strategies were successfully employed: (1) Differential Display (DD) (Liang, P., et al. 1994; Bauer, D., et al. 1993), (2) Representational Difference Analysis (RDA) (Braun, B.S., et al. 1995; Hubank, M., and Schatz, D.G. 1994; Diatchenko, L., et al. 1996) and (3) standard subtractive hybridization (Li, W.-B., et al. 1994; Harrison, S.M., et al. 1995). The latter underlies the present invention. A key feature is that the differentially expressed cDNAs have a high probability of being full-length. This facilitates a rapid transition to functional studies. The two former techniques were utilized because of "visual" nature (DD) and the ability to generate representative, differentially expressed probe populations (RDA) in a rapid manner.

Murine stem cell gene expression profiles. As a first step, a series of high quality, representative cDNA libraries were generated. The cDNA populations were directionally cloned into the pSport-1 or pSport-2 plasmids (BRL-Gibco). The most important libraries originate from purified stem cells. In one case, enough AA4.1+Lin-/loSca+ckit+cells were purified to allow construction of a non-based library using

standard methodologies. This library contains $\sim 4 \times 10^6$ independent recombinants (average cDNA insert size of 1-2 kilobases, kb). A second library was constructed using a new PCR-based technology called cap-finder (Clontech) designed to yield full-length cDNA copies. cDNAs ranging from 1 to 4 kb were commonly attained using this technique. For cap-finder procedures the purified cells were processed into DNase-I digested, poly-A⁺ mRNA according to microscale procedures routinely in use. The synthesis of cDNA was done with an aliquot of mRNA corresponding to approximately 20,000 cell equivalents. It has not been necessary to use less material, therefore this is not a lower limit. A Not I restriction site was included in the 3' reverse transcriptase primer (cap-finder, version 2) to facilitate directional cloning. An aliquot of the cDNA was amplified for varying PCR cycle numbers, and analyzed by Southern blots (pseudo-Northerns) for the presence of full-length copies of b-actin, GAPDH, CD18, flk2/flt3, cdk4, CD34 and other mRNAs. Optimal cycle numbers were used to amplify the remaining cDNA. The cDNA was cloned into the pSport-1 plasmid. The AA4.1+Lin-
15 /loSca+ckit+ cap-finder library contains $\sim 3 \times 10^6$ independent recombinants. Single-pass sequence analysis of random clones from this library indicated that 95% were full-length (based on sequences with an exact match in Genbank). Much more extensive sequencing of numerous clones from subtracted libraries has confirmed this. Other libraries constructed in similar ways include two libraries from AA4.1 cells, and two libraries from
20 AA4.1+Lin-/loSca-ckit+ cells. Finally, a library was constructed from AA4.1+Lin-/loSca+ckit+ cells cultured for 5-7 days in a differentiation promoting cytokine cocktail (IL3, IL6, KL). Competitive repopulation and in vitro progenitor assays confirmed a complete loss of LTRA with a significant retention of progenitor cells. All of the cDNA libraries are large ($> 2 \times 10^6$ independent recombinants) and of high quality (1-2 kb average insert size). In summary, the most primitive, intermediate and most mature
25 members of the hematopoietic hierarchy were "converted" into representative panels of expressed genes.

Subtractive Hybridization. These cDNA libraries were used in subtractive hybridizations to enrich for differentially expressed genes. Target libraries from
30 AA4.1+Lin-/loSca+ckit+ cells were subtracted with an AA4.1 driver cDNA library. This yields a population of cDNA clones which is enriched in sequences expressed in primitive stem/progenitor cells but not in mature cell types. In practice, a single-stranded target library is hybridized to an excess of in vitro synthesized biotinylated RNA from the

driver library as described . The opposite orientations of the cloning sites in pSport-1 and 2 insure target and driver complementarity. The target library was enzymatically converted to a single-stranded form using gene II protein and exonuclease III (Gene-Trapper protocols, BRL-Gibco). Previously, this was done by infection with M13 helper
5 phage often resulting in a bias for small cDNA inserts. Here, driver/target combinations were subtracted two times in order to facilitate the removal of commonly expressed sequences (verified by elimination of "housekeeping" genes such as b-actin and GAPDH). Concomitant enrichment of known, differentially expressed genes is also verified. Generally flk2/flt3 and CD34 probes were used. Both are expressed in the
10 AA4.1+Lin-/loSca+ckit+ subpopulation but not in the AA4.1 population. Following subtraction, the relative number of clones is reduced by up to 200-fold. In some cases (subtractions with AA4.1 material) the number of clones which "survive" the subtraction is on the order of 10-20,000. Because individual sequences may be represented more than one time, this does not necessarily imply that there are 10-20,000 differentially
15 expressed genes. The exact number of unique sequences (complexity) in the pool of subtracted clones must be determined. A more thorough discussion of complexity is found in a subsequent section. These subtracted libraries should be enriched for sequences expressed in the primitive portion of the stem/progenitor cell hierarchy; that is in stem cells and/or in primitive clonogenic progenitor cells. Two other subtracted
20 libraries, potentially enriched for sequences expressed in the most primitive stem cell but not in clonogenic progenitors were derived by subtracting the AA4.1+Lin-/loSca+ckit+ libraries with material from the closely related AA4.1+Lin-/loSca-ckit+ subpopulation.

Each subtracted library is arrayed at high density onto nylon membranes. Each clone in the array has a unique address in microtiter plates. A density of 20-30,000 clones
25 on a 22 x 22 cm. membrane is practical. Because of the inherent "noise" in any subtraction scheme, a positive selection criterion may be imposed in order to focus on true differentially expressed sequences. Details are presented in a subsequent section. To analyze the subtracted libraries, a high throughput sequencing effort was employed using three libraries. These are: 1) two AA4.1+Lin-/loSca+ckit+ cell libraries (standard and
30 cap-finder) subtracted extensively with AA4.1 cell material and 2) a standard AA4.1+Lin-/loSca+ckit+ library subtracted with AA4.1+Lin-/loSca-ckit+ material. An average of 400 bp of 5' sequence was obtained from about 1000 clones. To facilitate a rational handling of sequence information and to focus attention on a small number of

clones for functional analysis a "flow of information" strategy was devised. (See, Figure 1).

A major component of this strategy is bioinformatics. This can be global (comparisons with outside databases) and local ("in-house" analyses within an individual library or comparison of several libraries). A relevant example of the latter is comparison of genes from murine stem cells with genes derived from their human counterparts. Global bioinformatic analysis provides much information. First, it establishes if a given sequence corresponds identically or closely to an already identified gene in the mouse, human or other mammalian species. Such homologies can be detected at the nucleotide level. This can provide evidence of hematopoietic expression for a previously described gene. Second, where the homologies are statistically significant but not identical, novel members of gene/protein families can be identified. Third, a wider net can be cast over the databases by using conceptual translation of a sequence in the homology comparisons. The Examples provided herein illustrate the power of this approach in revealing similarities to proteins from invertebrate organisms such as *Drosophila*, *C. elegans*, and even yeast. In many cases these proteins have functions which have been uncovered by the analysis of mutants. Based on protein homologies "virtual links" are drawn between developmental regulation in invertebrates (such as in germ-line development) and in hematopoietic stem cells. The Notch/Notch ligand pathway first defined in invertebrate cell-fate determination and recently implicated in hematopoietic regulation is a good example.

A number of other putative proteins which share homology with key *Drosophila* proteins have been identified. Bioinformatics facilitates the recognition of peptide motifs such as EGF-like repeats, Ig-like domains or Zn-finger modules. Fourth, because the databases are annotated, predicted protein sequences can be assigned to cellular processes such as signal transduction pathways or apoptosis. It is also possible to categorize clones according to potential involvement in other mammalian stem cell systems such as the intestine. Fifth, it is feasible to perform virtual expression studies and to construct overlapping EST contigs which can yield virtual full-length cDNAs.

The following discussion summarizes the general murine findings and highlights a panel of "interesting" subtracted clones. A number of full-length sequences have been determined. Bioinformatic analysis summaries on a collection of 863 clones derived

from AA4.1+Lin-/loSca+ckit+ libraries subtracted with AA4.1 RNA are shown in Table 1.

The present invention encompasses numerous bioinformatic search and comparison parameters. This is the first analysis of its kind in the hematopoietic system and several important points emerge. First there is a high proportion (~50%) of novel sequences. Sixteen percent do not show homologies in any EST database queried. A preliminary analysis of a small sequence set from an AA4.1+Lin-/loSca+ckit+ library depleted of sequences in common with the AA4.1+Lin-/loSca-ckit+subset has indicated an even higher percentage of novel genes. These are higher percentages than would emerge from a random analysis of an unsubtracted library. This was directly addressed by performing these analyses on a similar number of sequences selected at random from one of the mouse EST databases. Second, among the ~50% of the clones which are significantly homologous to previously identified genes or proteins, the proportion of homologies to "housekeeping" genes is low. This underscores the effectiveness of the subtraction strategy. Third, an internal analysis of the data base has revealed few redundancies. This illustrates the degree of gene expression diversity between the two endpoints of the mouse hematopoietic hierarchy; thus supporting the starting hypothesis. A recent report has suggested that many sequences in mature blood cells are expressed in primitive, non-committed cells (Hu, M., et al. 1997). The data indicate that many of these are removed by the subtraction; thus uncovering a large, previously not described set of genes. Fourth the sequences with homologies to known genes or proteins can be subdivided according to protein families and putative function. Interestingly, a large percentage (32%) fall into the signaling protein category. Examples of these are described herein below.

As part of a more sophisticated bioinformatics approach, automation of database searches, information cross-referencing, and annotation was employed. An illustrative example is automated weekly database queries with the sequence set. New, previously-unidentified homologies were automatically noted and reported. In the mouse studies several genes were encountered whose expression would be predicted to differ between the AA4.1+Lin-/loSca+ckit+ and the AA4.1- populations. Both murine CD34 and flk2/flt3 were identified in the sequenced population. This provides a good internal control for the screening strategy. As shown in Table 2, a "short list" of identified

molecules was generated based on bioinformatics and in many cases, expression-specificity verification. The features of some of the molecules merit discussion.

The SA49P1 clone is homologous to sex comb on midleg; a member of the *Drosophila* Polycomb group of zinc-finger transcriptional repressors (Bornemann, D., et al. 1996). Polycomb proteins are regulators of homeobox (HOX) genes and maintain the
5 developmental stability of transcriptional states (Simon, J. 1995). There is currently great interest in HOX gene function in the biology of hematopoietic stem cells (Sauvageau, G., et al. 1995; Lawrence, H.J., et al. 1996). The C4-80 gene is homologous to the *Drosophila* cornichon gene (Roth, S., et al. 1995). Cornichon is required during
10 oogenesis for the induction of follicle cells, which provide the environment that supports oocyte development. Cornichon is a component of the *Drosophila* EGFR signaling pathway (Neuman-Silverberg, F.S., and T. Schupback 1996). The LL2-12 gene is similar to the *Drosophila* neurogenesis gene brainiac. Mutants defective in this extracellular gene product show neural hyperplasia. Brainiac also plays a role in oogenesis (Goode, S., et al.
15 1992; Goode, S., et al. 1996). Mosaic experiments show that brainiac is required in the developing oocyte. Brainiac involvement in the *Drosophila* EGFR pathway has been suggested. The LL2-35 gene is similar to the *Drosophila* germ cell-less gene (Jongens, T.A., et al. 1992). The product of this gene is required for the specification of the *Drosophila* germ line. The ectopic expression of germ cell-less causes somatic cells to
20 adopt the characteristics of pole cells (destined for the germ line). The homologies of the above clones to *Drosophila* genes were identified at the predicted amino acid level. This underscores the utility of the bioinformatic approach. In all four cases the cDNAs represent novel unpublished murine genes. Clearly, the involvement of three of the *Drosophila* genes in the ultimate stem cell system, the germline, coupled with the
25 identification of homologous genes expressed in mouse hematopoietic stem /progenitor cells gives food for thought. Four other cDNAs identified are homologous to the *Drosophila* genes dishevelled and smoothened (mouse homologs are already described) as well as to kelch and discs large (not previously described in the mouse). DD116 was originally isolated in a DD comparison of AA4.1+ vs. AA4.1 cells. A cDNA clone
30 identified with the DD116 probe was sequenced; the predicted protein is homologous but not identical to the murine beige gene product. The beige mutation in the mouse causes bleeding, immune system disorders and a coat-color phenotype (Perou, C.M., et al. 1996; Fukai, K., et al. 1996; Barbosa, M.D., et al. 1996). Murine beige is thought to be the

homologue of a gene responsible for the Chediak-Higashi syndrome in human. This novel gene is likely to be the second member in the beige family; it is designated herein as "taupe". Two of the cDNAs in Table 2 were identified in a DD comparison of AA4.1+Lin-/loSca+ckit+ cells and cytokine differentiated, cultured cells. One focus was
5 molecules whose expression disappeared with the loss of stem cell activity in culture. Cyt28 and Cyt19 fulfilled this criterion and were used to isolate full length clones which have been completely sequenced. Cyt28 encodes a novel, seven-transmembrane domain receptor in the same family as mouse F4/80 and human CD97 (secretin receptor superfamily) (Hamann, J., et al. 1995; Baud, W., et al. 1995) . Cyt 19 is a novel putative
10 methyltransferase. In both of these cases protein family assignments were made after extensive bioinformatic analyses. At least in the case of Cyt28, an antibody will be a useful reagent because it may provide a means to further subdivide the stem/progenitor cell population. Such antibodies are provided by the present invention. The cDNA SA61 is similar to a newly discovered molecule called p62dok. p62dok is a tyrosine
15 phosphorylated protein which binds to rasGAP and is likely to be a common target for numerous tyrosine kinases including Abl and ckit (Carpino, N., et al. 1997; Yamanashi, Y., 1997) . Interestingly, p62dok is also constitutively phosphorylated in hematopoietic progenitor cells from chronic phase CML patients (Carpino, N., et al. 1997) . The predicted protein encoded by SA61 appears to be a second member of this protein class.
20 Other representative molecules are listed in Table 2. Two cDNAs were identified as homologous to putative apoptosis regulators (SBSA56 and LL5-68, two cDNAs homologous to genes translocated in leukemias (LL5-03 and B2-67), several homologies to putative chromatin proteins (C4-23, C3-25 and LL2-89) and a LIM-domain encoding cDNA (LL5-96). Most interestingly, three cDNAs (B3-77, C2-48, and LL2-76) are
25 homologous to genes expressed in intestinal crypts or during intestinal development.

Isolated cDNAs were hybridized to slot blots representing globally amplified cDNAs from pools of individual progenitor and mature cells. Because the developmental potentials of these cells have been measured (by replating of siblings) it is possible to graphically represent the expression of a gene in various stages of the hierarchy. Three
30 examples are shown in Figures 2A-2C. The Smc-34 cDNA is a completely novel sequence with a predicted leucine zipper and several potential membrane spanning domains. These examples underscore the value of "interfacing" the above-described approach with a single cell approach. Thus, it was demonstrated that a given gene is

expressed in at least multipotential progenitor cells and other members of the progenitor cell hierarchy.

Human stem cell gene expression profiles. It is also a goal of the present invention to identify genes expressed specifically in human stem cells. It is useful to use
5 a multi-species (mouse and human) approach to define stem/progenitor cell-specific gene profiles. One goal is to integrate mouse and human information. However, rather than proceeding directly to human homologs of the mouse genes already identified, an independent human effort was initiated. The rationale for this is multi-faceted. First, it is possible that some aspects of human and mouse stem cell biology are regulated in
10 different ways. Although it is likely that most regulatory pathways will be conserved, it must be kept in mind that many properties ascribed to the most primitive stem cell population have been rigorously proven only in the mouse. Clearly a human lymphoid-myeloid stem/progenitor cell exists. However, the exact degree of proliferative capacity (the ability to give rise to oligoclonal hematopoiesis) as well as the exact spectrum of
15 differentiation potentials of human stem cells have not yet been accurately measured. Several potential differences between mouse and human have already been mentioned. One additional indicator of differences may be the difficulty of gene-transfer into stem cells from large animals (Larochelle, A., et al. 1997). This is not likely to result solely from the quiescent status of the most primitive cells. Second, while it is usually possible
20 to find human homologs for individual mouse genes by manipulating hybridization stringency, such conditions will vary for different genes. Therefore, to find human homologs for a large pool of mouse genes (i.e. 100) may be more labor intensive and costly than to independently determine a sequence profile of differentially expressed genes from human stem/progenitor cells. Relationships to the murine panel can then be
25 determined electronically where it is easy to manipulate comparative parameters. Clearly, for some individual mouse genes it will be of great importance to physically identify human homologs. In some cases it may be possible to use existing human EST databases to quickly obtain the sequence of a human homolog. Third, the availability of a large panel of human sequences specifically expressed in stem/progenitor cells lends
30 itself to the application of various chip and array technologies. Such technologies will be instrumental in identifying which subsets of human stem cell specific genes are up or down-regulated in the highly clonogenic stem cells from diseases such as Acute Myelogenous Leukemia (AML) and other leukemias.- Highly purified human stem

cells for cDNA libraries were obtained from normal donor BM. CD34⁺ cells were obtained from the mononuclear fraction using an immuno-magnetic affinity device (Miltenyi MACS column) and stained with a mixture of monoclonal antibodies (mAbs) designed to identify both primitive and mature cell surface markers. To identify antigens on mature cells a lineage (Lin) cocktail of FITC conjugated mAbs was used. These mAbs were directed against: CD3, CD11b, CD15, CD19, CD20, glycophorin A, and CD71 (CD3, CD15, CD19 and CD20 from Becton Dickinson, CD11b and glycophorin A from Coulter, and CD71 from Ortho Diagnostics). To identify primitive cells mAbs were used which were directed against CD34 (biotin conjugated and detected via a streptavidin-allophycocyanin reagent; Coulter) and either CD38 (phycoerythrin conjugated, Becton Dickinson) or CD90 (Thy1) (phycoerythrin conjugated; Pharmingen). The stained cells were analyzed using a dual laser Becton Dickinson FACStar Plus flow cytometer. Cells of the desired phenotype were sorted into siliconized tubes. To obtain material from cultured cells, CD34⁺ enriched populations (isolated as described above), were first cultured in serum-free medium (IMDM, 5 mg/ml low density lipoprotein, 2 mM glutamine, human serum albumin, insulin, and transferrin,) supplemented with IL-3 (5ng/ml), KL (25ng/ml), and FL (25ng/ml). After 1-4 days of culture, the cells were purified as described above. Human libraries were constructed using the cap-finder technology, version 2. Amplified cDNAs originating from several independent BM donors were pooled. Representative cDNA libraries have been constructed from the following sources: (1) BM CD34⁺Lin⁻ cells (7.1X10⁵ independent recombinants), (2) BM CD34⁺Lin⁻CD38⁺ cells (1.9X10⁶ independent recombinants), (3) BM CD34⁻ cells (1.6X10⁶ independent recombinants), (4) CB CD34⁺ cells (4.3X10⁵ independent recombinants), and (5) CB CD34⁻ cells (2.9X10⁵ independent recombinants). The average cDNA insert size in all libraries is 1-2 kb. According to the comparative biological properties of the source material used for the mouse and human cDNA libraries, the following parallels can be drawn: (1) mouse AA4.1+Lin⁻/loSca+ckit⁺ \cong human CD34⁺Lin⁻, (2) mouse AA4.1+Lin⁻/loSca-ckit⁺ \cong human CD34⁺Lin⁻CD38⁺ and (3) mouse AA4.1 \cong human CD34⁻. Thus, in both species, collections of clones representing the beginning, the middle and the end-points of the hematopoietic hierarchy have been generated. Using procedures described above, the human libraries were subtracted and arrayed 3,000 clones from two libraries: (1) BM CD34⁺Lin⁻ subtracted

with BM CD34+Lin-CD38+ material (referred to as: 38-)38+) and (2) BM CD34+Lin-
subtracted with BM CD34- material (referred to as: 38-). Before initiating a sequencing
effort pilot studies were performed to improve the resolution of the screen. Specifically,
in order to develop techniques which would eliminate as much non-specific "noise" as
5 possible from the libraries. RDA (PCR-Select, Clontech) were used. There are several
salient features of RDA. First, it is rapid (2-3 days) and applicable to small cell numbers
(ie. several thousand). It does not require the generation of cloned cDNA libraries
therefore, numerous comparisons can be made simultaneously. Second, it yields short
cDNA fragments (generated by 4-cutter restriction enzymes) which are representative of
10 differentially expressed mRNAs. Third, it is based on the suppression PCR technique
which prevents overamplification of abundant messages (Diatchenko, L., et al. 1996). In
addition, sequences expressed in both populations in an RDA comparison do not amplify
exponentially. Differentially expressed populations of cDNA fragments are not obtained
by physical enrichment but rather by selective PCR amplification. Fourth and most
15 importantly, it simultaneously yields two populations each representing differentially
expressed genes in one of the two starting samples. This bi-directionality is valuable
because as discussed below it can simultaneously reveal up and down regulated genes.
RDA comparisons were made between CD34+Lin- and CD34+Lin-CD38+ and between
CD34+Lin- and CD34- cDNAs. The latter comparison was verified by hybridization. As
20 shown in Figure 3A (right panel), the control, non-subtracted RDA cDNA population (38-
) contains b-actin sequences which are missing in the two subtracted RDA populations,
38- and 38-)38+. A differentially expressed gene (HDD-2, described below) is enriched
in the 38- RDA population and at least retained in the 38-)38+ RDA population (Figure
3A, left panel). Two, bi-directional, RDA cDNA populations (38- and the converse 34-
25)38) were used to probe (See Figure 3) duplicate arrays of a subtracted 38- library (Figure
3B). The correct RDA probe hybridizes to considerably more clones than the incorrect
probe (compare the greater numbers of spots in the left as compared to the right panel).
Hybridization signals for each clone are doublets due to the arraying technique.
Hybridization signals with the incorrect probe (right panel) suggest that further
30 improvements to this strategy are worth pursuing. Because the subtractions are based on
different technologies, it was reasoned that clones which "survived" the library
subtraction and hybridized preferentially to the "correct" RDA probe would be more
likely to represent true differentially expressed genes. Preliminary data from the mouse

studies had already suggested that this was indeed the case. The arrayed clones were also hybridized to probe populations from the entire starting non-subtracted libraries in order to eliminate residual cDNA clones corresponding to abundant mRNA species.

In one analysis, a total of 77 clones (49 from the 38- subtraction and 28 from the -38-)38+ subtraction) which "obeyed" the RDA criteria were sequenced from the 5' end. In order to verify expression specificity, a series of pseudo-Northern blot analyses were performed. cDNA populations from independent purified sources of material was amplified. This is of particular importance with an outbred species such as humans. In these experiments additional material from CD34+Lin-CD90+ (Thy1+) (three separate purifications) and their CD90- (Thy1-) counterpart subpopulations was also included. Cytokine-cultured samples were also included. These were CD34+Lin- cells and their CD38+ counterparts purified after 1, 2 and 4 days of culture. The collective panel of cDNAs also includes four independently purified CD34+Lin- populations and their CD34+Lin-CD38+ counterparts. All of these amplified cDNAs have been arranged on numerous pseudo-Northern blots. In short, this allowed us to evaluate the expression of a given "sequence tag" in stem cell populations isolated by two different criteria, from a number of independent donors and after cytokine culture. Several interesting genes emerged from the sequence data set. Individual clones obtained from the -38-)38+ subtraction are designated 38 letter/number (i.e., 38A1) while those from the 38-)34+ subtraction are designated 34 letter/number (i.e., 34A1). One clone (38B5) was identified and determined to be human flk2/flt3. A second interesting clone (34B4) is closely related to a gene encoding TINUR (Figure 4A). Clone 34B4 may be a novel variant of TINUR due to a 25 amino acid, in-frame deletion. TINUR was identified as an orphan member of the steroid receptor superfamily (NGFI-B/nur77 subfamily)(Okabe, T., et al. 1995). TINUR has also been implicated in apoptosis. An additional clone (34F4) is highly homologous to DAP-kinase. This protein is a serine-threonine kinase which has been implicated in cytokine (IFN- γ -induced apoptosis (Deiss, L., et al. 1995). Clone 34F4 (DAP-Kinase) and 34B4 (TINUR) both exhibit a stem-cell restricted expression pattern. Clearly, the identification of two genes whose products are implicated in apoptosis and whose expression is largely restricted to human stem cells is of interest. In addition, a cDNA (34A5) was identified which is closely related to the MLF1 gene which is a translocation partner in t(3;5)(q25.1;q34)(Yoneda-Kato, N., et al. 1996). This translocation is associated with Myelodysplastic Syndrome (a stem cell disease) and

AML. The sequence homologies and restricted expression pattern of 34A5 is shown in Figures 4B and 4C. In Figures 4C and 4D (and also 5A, 5B, and 6B) there are twenty-one samples of capfinder-amplified cDNA from various hematopoietic populations. From left to right these are: four CD34+Lin- populations, three CD34+Lin-CD90+
5 populations, two CD34- populations, four CD34+Lin-CD38+ (obtained from the same BMs as the CD38- samples in lanes 1-4), two CD34+Lin- samples (obtained from the same BMs as the CD90+ samples in lanes 5 and 6), three CD34+Lin- populations obtained after 1, 2 or 4 days of culture and finally their three CD34+Lin-CD38+ counterparts. A recent study shows that, at least with some types of AML, the disease
10 can be transferred into NODSCID mice only by the leukemic CD34+CD38- subpopulation (Bonnet, D., and Dick, J.E. 1997) . These demonstrate MLF1 expression in normal stem cells. As shown in Figure 4D, nucleophosmin (NPM), the partner in this translocation was also identified. This is an example of non-specifically expressed "noise" in the screens. The 38G2 cDNA is closely homologous to the LTG9/MLLT3
15 gene located on 9q22 and involved in t(9;11) leukemia (Iida, S., et al. 1993) (Figure 4E). A degree of stem cell expression specificity has also been observed. All of the above cDNA clones are likely to be identical to the homologous, previously identified genes. However, the suggested involvement of DAP-kinase and TINUR in apoptosis necessitates their inclusion in any comprehensive consideration of stem cell apoptotic pathways.
20 Similarly, the expression of two genes associated with myeloid leukemias bears on any speculation regarding the primary transformation target cell as well as the origins of the ultimate clinical phenotype in these and other leukemic disorders.

Other known genes to were found that exhibit identity or very close homology including: (1) G0S3, a fos-related gene (Heximer, S., et al. 1996) (Figure 5A) and (2)
25 HLA-DR (Figure 5B). G0S3 shows a specific expression pattern, while HLA-DR expression appears to be more variable. The expression status of Class II MHC on the most primitive human BM stem cell population is not entirely clear. It has been suggested to be present in a primitive, multipotent progenitor population but not in the most primitive stem cells (Sutherland, H.S., et al. 1989; Verfaillie, C., et al. 1990) . If
30 true, this may suggest an additional negative selection parameter for future experiments designed to subdivide the stem/progenitor cell hierarchy. Of the first 75 sequences, 22 have no homologies in the databases or homologies only to ESTs. Expression analyses on these clones are in progress. One gene of particular interest is called HDD-2 as well as

34B5, 34E1 and 38A11. The three latter designations reflect its isolation from both the 38- as well as the 38-)38+ subtracted libraries. The designation HDD-2 reflects its independent isolation in a limited DD "first look" at molecular differences in the purified cell populations. The likely full-length sequence of HDD-2 is ~500 bp. It contains a short open reading frame of 89 amino acids (SEQ.ID.No.: 71). The predicted peptide sequence is shown in Figure 6A. The 3' cDNA sequence contains a poly-A tail preceded by the canonical AATAAA poly-adenylation signal. Neither the nucleotide nor the predicted protein sequences of HDD-2 show homologies in any known gene or EST database. The expression profile of HDD-2 shown in Figure 6B, demonstrates that it is stem cell restricted. Also shown below (Figure 6C) is HDD-2 hybridization to a dot blot with numerous human pA+ mRNA samples (Clontech). HDD-2 hybridization is only visible in kidney (the other "spots" are background). It was confirmed that HDD-2 corresponds to a single-copy human gene by genomic Southern blot (Figure 6D). In summary, the results of this very low throughput human sequencing effort substantiate the overall approach; that judicious pre-enrichments and selections will result in rapid identification of biologically interesting and often novel genes. Most importantly, these studies firmly establish the existence of genes whose expression correlates with the most primitive stem cell phenotype.--

Other data were generated using (1) a murine stromal cell line to support enriched human stem/progenitor cells, and (2) tetracycline regulatable retroviral expression vectors. The stromal cell line AFT024 is efficient in long-term, in vitro maintenance of LTRA in purified murine fetal liver and adult bone marrow populations (Terstappen, L.W.M.M., et al. 1991) . Additional data demonstrated highly-efficient retroviral-mediated gene-transfer into murine LTRA as well as into primitive in vitro progenitors during the AFT024 cocultures. Further, it has been shown that AFT024 is very effective in supporting ELTCIC. Specifically, the CD34+CD38- immunophenotype as well as the functional capacity of these cells is maintained. The latter was measured in limiting-dilution replating experiments. Moreover, it has been shown that limiting numbers of human CD34+Lin- cells can give rise to both B cells and NK-cells when cultured on AFT024. However, the supportive activities of AFT024 on mouse and human stem/progenitor cells have been indistinguishable. In addition, more than 500 sequences have been analyzed from an AFT024-specific subtracted cDNA library. A number of candidate stem cell regulators have been identified. Three of these are dlk (Moore, K.A.,

et al. 1997), a novel BMP/TGF- β superfamily member and a novel selectin-related molecule. The present invention contemplates identifying human stem/progenitor cell receptors and/or ligands for the AFT024 specific proteins.

In order to facilitate functional studies of stem cell gene products several retroviral gene-transfer vector systems were constructed and characterized. All of these employ the 293T cell retroviral packaging system (Kinsella, T., and Nolan, G. 1996). High titers of virus can be produced transiently without the time and labor consuming effort required for stable producer cell lines. Large cDNA populations can also be converted into virus populations (Kitamura, T., et al. 1995). It is preferable to have inducible (or repressible) vectors which are also selectable. Also, the single-transcription unit tetracycline (tet) repressible vector was modified (Hofmann, A., et al. 1996). This vector includes an enhancer/promoter deletion in the 3' LTR. The tet system is currently one of the best inducible expression systems available; regulation over a several hundred-fold range of expression is observed (Shokett, P.E., and Schatz, D.G. 1996). Cloning sites were introduced in order to insert cDNAs in a sense or antisense orientation. A green fluorescent protein (GFP) marker was also included in these vectors driven by a thymidine-kinase promoter (TK-GFP) (Yang, T., et al. 1996; Cheng, L., and Kain, S. 1996). The cDNA fragment of interest is under tet regulation while GFP is constitutively expressed in transduced cells. In order to confirm this, NIH3T3 cells are infected with a LacZ virus and GFP⁺ cells sorted. The GFP⁺ cells express LacZ in the absence of tet, while LacZ expression is undetectable in most cells after the addition of tet. The titer of this model construct is approximately 10⁴/ml. which is suitable for tissue culture studies.

Table 1. Known genes (or extensive aa homology): 436 (50.5%)

Grouped by function	
by nucleotide: 407	Signaling/receptors: 133
by amino acid only: 29	Translational/post-: 62
	Structural: 49
to mouse: 230	Transcriptional/post-: 45
to human/other: 206	Cell fate: 35
	Other: 87
	Unknown function: 25

Novel genes: 427 (49.5%)

Homologous only to expressed sequence tags: 288 (33.3%)
 No homology to any known nt or aa sequences: 139 (16.2%)

Table 2.

Putative family	Clone	Method	Cells compared	Notes
7-transmembrane receptor	Cyt-28	diff. display	SC* fresh vs. 7d-cultured	secretin R superfamily
Methyltransferase	Cyt-19	diff. display	SC fresh vs. 7d-cultured	
	C3-54	subtraction	SC vs. AA4'	
Aspartyl protease	SA7	subtraction	SC vs. AA4'	
Signal transduction molecules	SA61	subtraction	SC vs. AA4'	dok family member
	LL2-02	subtraction	SC vs. AA4'	
beige-related proteins	DD116	diff. display	AA4' vs. AA4'	BTK associated
Transcriptional regulators	SA49P1	subtraction	SC vs. AA4'	contains WD repeat
	LL5-96	subtraction	SC vs. AA4'	polycomb homol.
G-protein signaling	LL4-39	subtraction	SC vs. AA4'	includes LIM domain
	B1-66	subtraction	SC vs. AA4'	cell cycle transition
Apoptosis-related genes	SBSA56	subtraction	SC vs. AA4'	similar to BL34
	LL5-68	subtraction	SC vs. AA4'	SMT3A-related
Chromatin proteins	C4-23	subtraction	SC vs. AA4'	sim. to Requiem transcript. factor
	C3-25	subtraction	SC vs. AA4'	yeast HST2-like
	LL2-89	subtraction	SC vs. AA4'	NHP2-like
Other stem cells (intestinal crypts)	B3-77	subtraction	SC vs. AA4'	yeast SIS2-like
	C2-48	subtraction	SC vs. AA4'	homologous to A4 gene
	LL2-76	subtraction	SC vs. AA4'	homologous to C101
Genes involved in leukemogenic translocations	LL5-03	subtraction	SC vs. AA4'	homologous to EDPF
	B2-67	subtraction	SC vs. AA4'	t(1;11)(q21;q23)
Homologues of Drosophila developmental genes	LL2-35	subtraction	SC vs. AA4'	t(X;14)(q28;q11)
	LL2-12	subtraction	SC vs. AA4'	germ cell-less
	C4-80	subtraction	SC vs. AA4'	brainiac
	B4-14	subtraction	SC vs. AA4'	cornichon
				discs-large

*The designation "SC" refers to sorted day 14 fetal liver cells having the phenotype AA4.1⁺Lin⁻Sca⁺ckit⁻.

EXAMPLE 2: Molecular cloning and characterization of AA4, an early marker of hematopoietic development

In this example, the expression cloning and molecular characterization of AA4, a surface marker expressed on hematopoietic stem and progenitor cells is described. The results demonstrate that AA4 is a 130kDa type I glycosylated membrane protein whose structural organization suggests a role in cell adhesion. Expression analysis showed that high levels of AA4 are found in lung, heart, and bone marrow. It is not found in undifferentiated ES cells, but its expression is upregulated as these cells differentiate into colonies of hematopoietic precursors and endothelial cells. In the hematopoietic system, expression of AA4 correlates with the expression of stem cell markers CD34 and CD43. Functional studies revealed that AA4 coimmunoprecipitates with CD34 and CD43, suggesting that these proteins form a macromolecular complex which functions in the regulation of cell adhesion, proliferation and/or differentiation of hematopoietic cells.

Introduction

Monoclonal antibody AA4.1 was first described more than a decade ago (McKearn et al., 1984) and since then it has become a useful marker for the isolation and analysis of hematopoietic cells (McKearn et al., 1985; Jordan et al., 1990; Fujimoto et al., 1996). A number of works have shown that the antigen recognized by AA4.1 is present on a subset of primitive hematopoietic progenitors found at various stages of development in sites of active hematopoiesis in yolk sac (Cumano et al., 1993; Auerbach et al., 1996; Yoder et al., 1997), AGM region (Godin et al., 1995; Marcos et al., 1997), fetal liver (McKearn et al., 1985; Jordan et al., 1990; Cumano and Paige, 1993), and bone marrow (Li et al., 1996; Szivassy and Cory, 1993). In yolk sac, AA4-positive cells are first detected at day 8-10 of gestation (Cumano et al., 1993; Sanchez et al., 1996). At day 9-10 of gestation AA4⁺c-Kit⁺ progenitors are found in the P-Sp/AGM region (Sanchez et al., 1996; Marcos et al., 1997), and by day 14 of development, AA4 defines 0.5-1.0% of the fetal liver tissue that contains the entire hierarchy of primitive hematopoietic cells (Jordan et al., 1990). Proliferation within each successive compartment results in increased total number of progenitor cells. Antigen density per cell also increases with developmental progress, which is especially marked for c-Kit and AA4 (Marcos et al., 1997). In bone marrow, HSC are found in both AA4⁺ and AA4⁻ subpopulations, although in adult marrow AA4 is largely regarded to be a marker of early B lymphoid lineage. In addition,

recent studies demonstrate that the expression of AA4 parallels the onset of hematopoietic development in differentiating ES cells (Kabrun et al., 1997; Lin and Neben, 1997; Potocnik et al., 1997). Taken together, these results indicate that AA4 plays an important role in hematopoiesis and has to be studied in more detail.

5 In order to achieve this goal, AA4 was molecularly cloned and characterized. The results demonstrate that AA4 is a 130kDa type I transmembrane glycoprotein whose structure suggests a role in cell adhesion. Expression analysis showed that high levels of AA4 are found in lung, heart, and bone marrow. In the hematopoietic system, expression of AA4 correlates with the expression of stem cell markers CD34 and CD43. Functional
10 studies indicate that AA4 coimmunoprecipitates with CD34 and CD43, suggesting that these proteins form a macromolecular complex which may function in the regulation of cell adhesion, proliferation and/or differentiation of hematopoietic cells.

Results

Cloning of AA4. In order to identify AA4, AA4.1 monoclonal antibody was used
15 to screen various cell lines and primary hematopoietic cells. The murine B lymphoid cell line D2N was found to express relatively high levels of AA4 antigen (see Table 3). Immunoprecipitation of protein extracts prepared from D2N cells showed that AA4.1 recognizes a protein with apparent molecular weight (M_r) 130kDa (Figure 7). This protein was also present in extracts prepared from B lymphoid cell line M2.4, and AA4
20 positive hematopoietic cells derived from bone marrow and fetal liver.

 To isolate AA4, a cDNA library was prepared from D2N cells and cloned into the polylinker region of a retroviral expression vector REBNA (see Materials and Methods). Following production of retroviruses, NIH 3T3 cells were infected with the recombinant cDNA library and selected for AA4 expression by flow cytometry using PE-conjugated
25 AA4.1. After two rounds of sorting, genomic DNA extracted from AA4-positive cells was analyzed by PCR amplification using viral vector primers. This resulted in the amplification of a 3.1 kbp cDNA which was gel-purified and subcloned for further analysis. Infection of NIH 3T3 fibroblast or EML C1 hematopoietic cells with REBNA/AA4, a recombinant retrovirus expressing the cloned cDNA, has led to the
30 acquisition by cells of high affinity to AA4.1 mAb (Figure 8A). In infected cells, AA4.1 detects a 130kDa surface protein which comigrates with the endogenous AA4 from D2N cells (Figure 8B), thus indicating that the cloned cDNA encodes AA4.

Sequence analysis of AA4. Sequence analysis of the cloned cDNA (SEQ. ID No.: 72) showed that it has a single open reading frame encoding a protein of 644 amino acids (see Figure 9)(SEQ. ID No.: 73. The deduced amino acid sequence includes a putative leader peptide and the mature protein which starts at position 20. The protein
5 contains a long N-terminal extracellular region, a single putative hydrophobic transmembrane region, and 47 amino acids of the C-terminal cytoplasmic domain. The extracellular part of AA4 is composed of two major structural moieties. The N-terminal region contains a C-type lectin domain (CTL) which has 32% sequence homology to endothelial cell receptor thrombomodulin. This region is followed by a cysteine-rich
10 domain composed of six epidermal growth factor (EGF)-like repeats, three of which are consistent with the calcium-binding EGF motifs. Similar repeats are found in the extracellular domains of a large number of membrane-bound proteins and in proteins known to be secreted (Bork et al., 1996).

Database searches revealed that AA4 exhibits high homology to C1qRP, the
15 human receptor for complement component C1q expressed on surfaces of myeloid cell lineage and endothelial cells (Nepomuceno et al., 1997). Sequence alignment showed that AA4 and C1qR have approximately 68% identical amino acid positions and similar domain structures. Highest homologies were found within the N-terminal parts of the proteins and their C-terminal cytoplasmic domains (see Figure 10A and 10B), suggesting
20 that AA4 and C1qR may share functional properties.

The amino acid sequence of AA4 contains numerous potential O-linked and N-glycosylation sites. Although the predicted M_r of AA4 is 67.4 kDa, two protein bands exhibiting M_r 105kDa and 130kDa respectively, are immunoprecipitated by AA4.1 mAb
25 in cells infected with REBNA/AA4 (Figure 8C). When ^{35}S -labeled cells are chased with nonradioactive media, the intensity of the 105kDa band rapidly decreases, while the intensity of the 130kDa band increases, thus indicating that p130 is the mature form of the protein, whereas p105 is its precursor. In agreement with this conclusion, immunoprecipitation of biotinylated cells using AA4.1 reproducibly results in the detection of a 130kDa surface protein (see Figure 8B).

30 Expression patterns of AA4. Northern blot analysis showed that in adult mouse tissues, AA4 is expressed at high levels in lung, heart, and bone marrow. No detectable expression was found in brain, testis, spleen, and thymus (Figure 10B). In normal tissues

and transformed cell lines, a 7kb RNA species hybridizes with the cloned cDNA (Figure 11A and 11B). In addition to the 7kb species, poly(A)-RNA from D2N cells contains a minor band corresponding to a 3.2 kb mRNA (see Figure 11A, lane 8). Similarity search against expressed sequence tags (ESTs) showed that databases contain at least seven
5 different sequences corresponding to the 3'-untranslated region of the cloned *aa4*cDNA which were isolated from the following mouse tissues: colon (Genbank accession No AA929174), heart (AA435107), lymph node (AA185911 and AA267407), lung (AA220480), mammary gland (AI021507), and spleen (AA145088). Sequence identity with the ESTs abrogates upstream of nucleotide G at position 2481 in the 3'-untranslated
10 region of *aa4*cDNA, thus indicating that the cloned cDNA corresponds to an alternatively spliced 3.2 kb *aa4*mRNA.

Expression of AA4 in hematopoietic progenitor cells AA4 is produced in murine hematopoietic progenitor cells and immature B lymphocytes found at various stages of development in yolk sac (Godin et al., 1995; Marcos et al., 1997), fetal liver (McKearn et
15 al., 1985; Jordan et al., 1990), and bone marrow (Cumano et al., 1992; Li et al., 1996). To confirm that the cloned cDNA encodes AA4, RT-PCR was performed on hematopoietic cells fractionated using several different techniques. The cells analyzed included AA4⁺ and AA4⁻ fetal liver (FL) cells; AA4⁺ FL cells fractionated into Lin^{lo}c-Kit⁺Sca-1⁺ and Lin^{lo}c-Kit⁺Sca-1⁻ populations; and hematopoietic progenitors isolated from bone
20 marrow (BM) by cell sorting using combinations of different surface markers.

Figure 12A shows that *aa4* was amplified from AA4⁺ FL cells, whereas in AA4⁻ cells it was only present at low levels. AA4 expression was high in Lin^{lo}c-Kit⁺Sca-1⁺ cells enriched for HSC activity. In adult marrow cells, *aa4* was amplified from Lin⁺ cells and Lin⁻c-Kit⁺Sca-1⁺CD34⁺ multipotential progenitors. At lower levels *aa4* was present
25 in Lin⁻c-Kit⁺Sca-1⁺CD34⁻ long term reconstituting HSCs (Figure 12B). In all tested cells *aa4* expression correlated with the expression of stem cell markers CD34 and CD43.

Embryonic stem (ES) cells which have been shown to generate progenitors for most hematopoietic lineages during differentiation *in vitro* (Keller, 1995) were also tested in this experiment. Figure 12C shows that *aa4* was not found in undifferentiated ES cells
30 but its expression was upregulated as these cells differentiated and formed blast cell colonies (BL) and colonies of more differentiated hematopoietic cells (HMT). These results are in line with previous studies which showed that AA4 is expressed in ES-

derived hematopoietic precursors (Kabrun et al., 1997; Ling and Neben, 1997; Scott et al., 1997) and show that this expression is not abrogated upon differentiation of blast cell colonies into endothelial cells (ENT in Figure 12C).

Ectopic expression of AA4 has no mitogenic effect EML C1 and NIH 3T3 cells
5 expressing exogenous AA4 did not exhibit morphological changes indicative of alterations in their growth properties. To examine the effect of AA4 on growth phenotypes in more detail, NIH 3T3 and primary mouse embryo fibroblasts were infected with a retrovirus expressing AA4 and maintained in high and low serum conditions. FACS analysis of the transduced cells confirmed that in each case the efficiency of
10 infection was close to 100%. However, examination of growth rates showed that overexpressed AA4 had no apparent effect on proliferation of both cell types as compared to control uninfected cells or the corresponding cells infected with a retrovirus expressing GFP.

AA4 coimmunoprecipitates with CD34 and CD43 To investigate interactions
15 with other proteins, hematopoietic and fibroblast cells expressing AA4 were immunoprecipitated with AA4.1 mAb and then examined by immunoblot analysis using a panel of antibodies directed against membrane-associated proteins. This analysis revealed that AA4 coimmunoprecipitates with CD34, a membrane glycoprotein selectively expressed within the hematopoietic system on stem and progenitor cells, and CD43 which
20 is a major O-glycosylated sialomucin found on the surfaces of most leukocytes. This result is in line with previous studies which showed that CD43 coimmunoprecipitates with the human C1qRP (Guan et al., 1991; 1994). Figure 13B shows that in the murine D2N and EML C1 cells, a 52kDa protein is the major isoform of CD43 that associates with AA4. This 52kDa protein was found to be reactive with both the N-terminal (S19)
25 and C-terminal (M19) anti-CD43 antibodies, indicating that it is not a breakdown product. In NIH 3T3 fibroblasts coexpressing CD43 and AA4, a 54kDa and a 170kDa CD43 isoforms coimmunoprecipitated with AA4 pointing to the glycosylation differences between CD43 expressed in different cell types. A 115kDa CD43 isoform which
30 previously have been shown to be sialylated and thus overly negatively charged (Guan et al., 1994) did not form macromolecular complexes with AA4 in NIH 3T3 cells. Similarly, AA4 did not coimmunoprecipitate with a 115kDa CD43 isoform found in D2N cells and a 120kDa isoform found in EML C1 (Figure 13B).

To examine associations of AA4 with CD34, both proteins were expressed in Rat-1 cells following infection with the corresponding retroviruses. See Figure 13C.

In Rat-1 cells coexpressing AA4 and CD34, AA4.1 the results also demonstrate that AA4 coimmunoprecipitates with CD34. Studies have suggested a role for both CD34 and CD43 in the regulation of adhesion, growth and differentiation of hematopoietic precursors (Ardman et al., 1992; Bazil et al., 1997; Chen et al., 1996; Suzuki et al., 1996; Wood et al., 1997; Stockton et al., 1998). Coexpression of AA-Fc fusion protein in NIH 3T3 cells stably producing AA4 showed that AA4-Fc forms heterodimeric complexes with AA4. Figure 11A shows that equimolar amounts of AA4 and AA4-Fc were coprecipitated from these cells by protein A, indicating that AA4 is prone to homo- or heterodimerization.

Discussion

This Example describes the expression cloning and sequence analysis of AA4, a molecular marker expressed on hematopoietic stem and progenitor cells. The cDNA encoding AA4 was isolated from a retroviral cDNA library prepared from the murine D2N lymphoid cell line. Sequence analysis of the cloned cDNA revealed that AA4 is a type I transmembrane protein composed of 625 amino acids. The extracellular part of the molecule contains two major structural moieties, a C-type lectin carbohydrate recognition domain and six EGF-like domains. Similar repeats have been found in a large number of membrane-bound proteins or in proteins known to be secreted. The cytoplasmic domain of AA4, in contrast, bears no structural similarity with known protein families. Instead, AA4 revealed strong homology to ClqR, the human receptor for complement component Clq which is predominantly expressed in phagocytic cells such as monocytes, neutrophils, and endothelial cells (Nepomuceno et al., 1997). Sequence alignment shows that AA4 and ClqR have approximately 68% identical amino acid positions and share similar domain structure. Highest homologies were found within the Nterminal parts of the two proteins and their C-terminal cytoplasmic domains, suggesting that AA4 and ClqR may have similar functional properties. Expression of AA4 correlates with the expression of CD34 and CD43, two other surface markers normally present on murine hematopoietic stem and progenitor cells. Previous studies have shown that AA4-positive cells are first detected in yolk sac at day 8-10 of gestation (Cumano et al., 1993; Sanchez et al., 1996).

At day 9-10 pc AA4+c-Kit+Mac-l+ progenitors are found in the P-Sp/AGM region (Sanchez et al., 1996; Marcos et al., 1997), and by day 14 of gestation, AA4

defines 0.5-1.0% of the fetal liver tissue that contains the entire hierarchy of primitive hematopoietic cells (Jordan et al., 1990). As proliferation within each successive compartment results in increased total number progenitor cells, antigen density per cell also increases with developmental progress, which is especially marked for c-Kit and AA4 (Marcos et al., 1997). A significant proportion of hematopoietic progenitors from yolk sac also express CD34, CD38, CD43, CD44, and Mac-1 however show little or no expression of CD4, CD8, CD45R and Sca-. These same markers are present on FL-derived HSC which also express Sca-1. Adult marrow HSCs also express cKit, CD38 and Sca-1, but do not normally express Mac-I or AA4. In BM, HSC are found in both AA4+ and AA4-subpopulations, although in adult marrow AA4 is largely regarded to be a marker of early B lymphoid lineage 0. These results indicate that the expression of cell surface antigens changes on HSC during ontogeny and that differential display of these cell surface markers may reflect relationship between HSCs that contribute to multilineage hematopoiesis and distinct anatomical sites during development.

15 **Methods**

Cells and tissue culture. NIH 3T3 fibroblasts were grown in DME medium supplemented with 10% fetal calf serum (FCS). D2N cells were grown in RPMI 1640 medium containing 10% FCS. EML C1 cells were grown in IMDM supplemented with 20% horse serum and 8% BHK/MKL conditioned medium (Tsai et al., 1994). To maintain the multipotentiality of EML C1, the cells were kept at low density (0.5 - 5x10⁵/ml) and subcultured every two days. Cell lines constitutively expressing AA4 and GFP were derived from NIH 3T3 or EML C1 cells by infecting with the corresponding recombinant retroviruses.

Retroviral-mediated gene transfer. Retrovirus expression vector REBNA was constructed by substituting the LacZ gene contained within the EcoRI-NotI fragment of plasmid LZRSPBMN-Z (Kinsella and Nolan, 1996) with a synthetic polylinker composed of EcoRI, XhoI, SfiI, and NotI sites. Retroviral vector REBNA/IRESGFP contains a poliovirus IRES element and the CDNA encoding color-enhanced GFP (S. Zolotukhin, Gainesville, FL) inserted into REBNA.

For DNA transfections, cells were plated at a density 2x10⁶ cells per 60mm dish and transfected with 5ug of plasmid DNA using 20ul of lipofectamine reagent (GibcoBRL). REBNA-transfected cells were selected in puromycin (2ug/ml) and grown to confluence prior to collecting virus supernatant. For infections" the culture medium

was supplemented with polybrene (Sigma) at 5ug/ml. The appropriate virus was added and incubated overnight. Multiple *infections were carried* sequentially, with the *appropriate* selection between.

Isolation of RNA and CDNA cloning. Poly(A) RNA prepared from D2N cells
5 was converted into CDNA using Superscript II Reverse Transcriptase (GibcoBRL) and an
oligo(dt) primer containing NotI site, 5'-TGGTGTCTGACGCAGAGTAGCGGCCGCT18
(SEQ.ID.No.:74). The second strand was synthesized using DNA polymerase I in
combination with *E. coli* RNase H and *E. coli* DNA ligase as described (Gubler and
Hoffman, 1983). An adaptor composed of complementary oligonucleotides, 5'-
10 GGCCCGGGCCGGCC (SEQ.ID.No.:75) and 5'-TCGAGGCCGCCCCGGGCC
(SEQ.ID.No.:76), was ligated to the CDNA and cut with NotI to produce CDNA
molecules with NotI and XhoI termini for directional cloning. After size fractionation in
agarose gel, cDNAs larger than 2.5 kbp were ligated into NotI and XhoI cut plasmid
REBNA and electroporated into electrocompetent DH12S cells (GibcoBRL). Plasmid
15 DNAs were transfected into 293-derived packaging cell line for retrovirus production.
Virus-containing supernatants were collected and stored at -80C.

NIH 3T3 cells infected with the recombinant retroviruses representative of D2N
CDNA library were selected for the production of AA4 by flow cytometry using
phycoerythrin-conjugated AA4.1 mAb. After two rounds of sorting, genomic DNA
20 isolated from AA4-positive cells was subjected to pcr amplification using retroviral
vector primers, 5'CAGCCCTCACTCCTTCTC (SEQ.ID.No.: 77) and 5'-
GGTGGGGTCTTTCATTCC (SEQ.ID.No.: 78) (Kitamura et al., 1995). Amplified
CDNA was gel purified and subcloned into pbluescript SK and REBNA plasmid vectors.
Nucleotide sequences were analyzed using NCBI Blast database search programs and
25 ExPASy molecular biology server from the Swiss Institute of Bioinformatics.

Northern blot hybridization. RNAs prepared using acid guanidinium thiocyanate-
phenol extraction procedure (Chomczynski and Sacchi, 1987) were separated on
formaldehydeagarose gels and blotted onto the Hybond-N nylon membranes
(Amersham). Hybridization probes were derived from cloned cDNAs using Ready To
30 Go DNA labeling beads (Pharmacia Biotech). Hybridizations were performed as
described previously (Petrenko et al., 1997).

Cell labeling and Immunoprecipitations. The ECL protein biotinylation system
(Amersham) for the detection of cell surface proteins was used as recommended by the

manufacturer. For radioactive labeling, 2×10^6 cells were incubated with ^{200}OuC i Translabel (ICN) in 2ml of cysteine and methionine-deficient medium for 2hr at 37°C . Cells were washed in PBS and lysed 10min. on ice in 500ul of NP40 buffer containing 20mM TrisHCl pH 7.6; 150mM NaCl; 0.5% NP40; 1mM PMSF; 5mM benzamidine; 5 1mM sodium vanadate; 10ug/ml aprotinin; 10ug/ml leupeptin. Lysates were cleared by centrifugation and 2ug of soluble antibody were added to the supernatants followed by 25ul of Protein G-Sepharose suspension (Sigma). After 2 to 4hr incubation at 40°C with rotation, protein G-antibody complexes were pelleted and washed in the successive changes of wash buffer I (1M NaCl; 10mM TrisHCl pH 8; 0.1% NP40); wash buffer 2 10 (0.1M NaCl; 10mM TrisHCl pH 8; 0.1% NP40); wash buffer 3 (10mM TrisHCl pH 8; 0.1% NP40). Samples were eluted by boiling 2min. in 2Xgel loading buffer, separated by SDS/PAGE, dried and exposed to X-ray film.

Western blot analysis. Protein extracts for Western blot analysis were prepared as described (Morrison et al., 1991). The antibodies used included goat anti-mouse CD43 15 polyclonal IgG (M19 and S-19, Santa Cruz) in combination with HRP-conjugated secondary antibodies and ECL detection system (Amersham).

Flow cytometry and RT-PCR. Timed-pregnant mice and 5- to 7-week-old female mice (C57Bl/6j) were purchased from Jackson Laboratory (Bar Harbor, ME). AA4-positive cells were isolated from day 14 fetal liver by immunopanning on Petri dishes 20 coated with AA4.1 antibody (10ug/ml). Hematopoietic stem cells were purified from AA4-positive fraction by staining with lineage-specific antibodies as described previously (Moore et al., 1997). Three-color fluorescence-activated cell sorting for lineage negative to low, Sca-1(+), c-Kit(+) cells was performed on a multilaser FACS Vantage with CellQuest software (Beckton Dickinson). ES cells differentiated into blast cell colonies, 25 hematopoietic progenitors, and endothelial cells prepared as described previously.

For RT-pcr, poly(A)-RNA isolated from sorted cells was converted into cDNA using Superscript II Reverse Transcriptase (GibcoBRL) and CapFinder cDNA amplification kit (Clontech). Gene-specific primers for pcr amplification included: 5'-TTCAGCAAGCCCTGACTC (SEQ.ID.No.:79) and 5'-GCCACCTTCGAAGCAATC 30 (SEQ.ID.No.:80) (AA4); 5'-GAGCGGTACAGGAGAATG (SEQ.ID.No.:81) and 5'-GCCCCACCAACCAAATCA (SEQ.ID.No.:82) (CD34); 5'-ACCGCGTTCTTCTGTAAC (SEQ.ID.No.:83) and 5'-CAGCTAACAGCAGGATCC

(SEQ.ID.No.:84) (CD43); G3PDH Control Amplimer Set (Clontech) for the amplification of GAPDH.

**EXAMPLE 3: *IN VITRO* MAINTENANCE OF HIGHLY PURIFIED,
TRANSPLANTABLE HEMATOPOIETIC STEM CELLS**

The cellular and molecular mechanisms which regulate even the most primitive
5 hematopoietic stem cell are not well understood. This example details a systematic
dissection of the complex hematopoietic microenvironment in order to define some of these
mechanisms. An extensive panel of immortalized stromal cell lines from murine fetal liver
was established and characterized. Collectively, these cell lines display extensive
heterogeneity in their *in vitro* hematopoietic supportive capacities. This example describes a
10 long-term *in vitro* culture system, utilizing a single, stromal cell clone (AFT024) that
qualitatively and quantitatively supports transplantable stem cell activity present in highly
purified populations. Disclosed is multi-lineage reconstitution in mice that received the
equivalent of as few as 100 purified bone marrow and fetal liver stem cells which were
cultured for 4-7 weeks on AFT024. The cultured stem cells meet all functional criteria
15 currently ascribed to the most primitive stem cell population. The levels of stem cell activity
present after 5 weeks of coculture with AFT024 far exceed those present in short-term
cytokine-supported cultures. In addition, the maintenance of input levels of transplantable
stem cell activity is accompanied by the expansion of other classes of stem/progenitor cells.
This suggests that the stem/progenitor cell population is actively proliferating in culture and
20 that the AFT024 cell line provides a milieu which stimulates progenitor cell proliferation
while maintaining *in vivo* repopulating activity.

Introduction

Mammalian blood formation originates in a small population of hematopoietic stem
cells. The hallmark features of these cells are: (1) a hierarchical multilineage differentiation
25 potential with the ability to clonally give rise to at least 8 distinct cell lineages, (2) self-
renewal capacity which is reflected in the life long continuous activity of few, in some cases
single, stem cells and (3) a dramatic proliferative potential which is ultimately responsible
for the production of large numbers of mature blood cells. (Lemischka, I.R. 1992;
Morrison, S.J. et al. 1995; Harrison, D.E. 1980). During the past decade much progress has
30 been made in providing a physical phenotype for this rare population of stem
cells.(Spangrude, G.J., et al. 1995; Lemischka, I.R. 1992; Bauman, J.G., et al. 1998).
However, currently, the only reliable functional assay system for the most primitive stem
cell compartment is long-term *in vivo* transplantation. No *in vitro* system has been

developed which adequately recapitulates stem cell behaviors. Therefore, the cellular and molecular mechanisms that regulate the biology of stem cells have remained obscure.

A major challenge in stem cell research is the establishment of culture systems which facilitate *in vitro* maintenance of long-term transplantable stem cell activity. This is a necessary first step towards a cellular and molecular understanding of the regulatory mechanisms which mediate commitment versus self-renewal decisions. Moreover, the establishment of such culture systems is a prerequisite for the potential expansion of undifferentiated stem cell populations as well as for the generation of stem/progenitor cells committed to selected lineages.

Efforts to develop culture systems for the maintenance of transplantable stem cells can be subdivided into two broad categories: (1) those utilizing defined cytokine combinations as the only culture supplements and (2) those relying on a pre-established stromal monolayer as an additional supportive component (with or without exogenously added cytokines). Both of these strategies have met with only limited success. In the first case, it has been repeatedly demonstrated that combinations of cytokines can exert potent stimulatory effects on stem/progenitor populations. (Spangrude, G.J., et al. 1988; Jordan, C.R. et al. 1990; Fleming, W.H., et al. 1993). In some studies, highly purified stem cells (Li, C.L., and Johnson, G.R. 1990; Spangrude, G.J. and Johnson, G.R. 1990). have been used and the direct effects of cytokines have been demonstrated at the single cell level. (Jones, R. et al. 1990; Uchida, N. et al. 1993). While informative, the vast majority of these studies are limited by their strictly *in vitro* nature. Thus, it is feasible to expand, replatable *in vitro* progenitor populations (Li, C.L., and Johnson G.R. 1992; Uchida, N. et al. 1993) and to stimulate colony-formation by cells with both myeloid-erythroid and lymphoid potentials, Jones, R., et al. 1990; Uchida, N. et al. 1993; Trevisan, M., and Iscove, N.N. 1995; Ogawa, M. 1993) however, the equivalence of these progenitor cells with the *in vivo* transplantable stem cell population remains speculative. Several studies have clearly demonstrated a dramatic loss of *in vivo* repopulating potential as a result of cytokine driven *in vitro* proliferation. (Knobel, et al., 1994; Peters, et al., 1995; Traycoff, et al., 1996). A small number of studies have shown that defined cytokine combinations promote the maintenance of transplantable activity. (Rebel., et al., 1994). However, most of these are limited both by the use of very short culture periods, the exact nature of the *in vivo* assay, and the use of non-enriched stem cell sources. (Muench, et al., 1993; Holyoake, et al., 1996; Soma, et al.,

1996). This precludes interpretations suggesting a direct action of the given cytokine(s) in maintaining transplantable activity.

A further complication with defined cytokine studies is the inability to ascribe *in vivo* physiological relevance to the observed effects. It has long been accepted that in the intact animal, stem cells are found in close association with discrete cellular microenvironments.(Lord, et al., 1975; Trentin, et al., 1970; Weiss, et al., 1991; Wolf, 1979). These observations suggest both the existence of stem cell niches and the notion that *in vivo* stem cell regulatory mechanisms are likely to require cell-cell contact or short range interactions.(Dorschking, 1990). Efforts to understand the features of the hematopoietic microenvironment began with the establishment of the Dexter long-term culture (LTC) system. (Dexter, et al., 1977). In this culture system hematopoiesis is maintained for weeks or months by a heterogeneous adherent cell monolayer derived from bone marrow (BM). While some degree of transplantable stem cell maintenance and self renewal (Fraser, et al., 1990) has been demonstrated, a general feature of the Dexter-LTC is a dramatic net decrease of stem cell activity over time. (Harrison, et al., 1987; Van der Sluijs, et al., 1993). Although much progress has been made, especially in studies of human stem/progenitor cells, (Sutherland, et al., 1989; Verfaillie, et al., 1995; Hao, et al., 1996) a further drawback of this system is the heterogeneity of the stromal cell types present in the supportive monolayer. This hampers the identification of regulatory mechanisms. Studies have been reported where the heterogeneous stromal monolayer is replaced with cloned stromal cell lines. (Roberts, et al., 1987; Kodama, et al., 1984; Issad, et al., 1993; Wineman, et al., 1993). Many of these cell lines can support *in vitro* myelopoiesis, (Suzuki, et al., 1992; Neben, et al., 1993; Kodama, et al., 1992) B-lymphopoiesis (Collins, et al., 1987; Whitlock, et al., 1982) or in some cases both. Wineman, et al., 1993; Wineman, et al., 1996) However, very few studies have focused on the *in vitro* maintenance of the most primitive transplantable stem cell compartment. Moreover, with one exception (Szilvassy, et al., 1996) the studies which have focused on the *in vitro* maintenance of this stem cell population begin with heterogeneous unpurified sources of hematopoietic activity. Wineman, et al., 1992, 1996; Deryugina, et al., 1994). Such populations contain numerous non-hematopoietic stromal cell types. Therefore, it has not yet been possible to assign a direct stem cell supporting phenotype to a given stromal cell line.

It was hypothesized that the rare frequency of primitive, stem cells may suggest an equally rare frequency of stem cell supporting microenvironmental niches. Accordingly, we

established and characterized a large panel of conditionally immortalized, cloned stromal cell lines from mid-gestation fetal liver. This organ was chosen because, during development, it is here that the stem cell compartment is undergoing self-renewal expansion in addition to differentiation. (Moore, et al., 1970). The cell lines were generated as previously described, (Wineman, et al., 1996) by immortalization with a temperature sensitive SV40TAg. (Frederiksen, et al., 1988) The clonal nature of the AFT024, 2018, and 2012 cell lines was verified by Southern blot analysis which detected a single, unique proviral integration locus in their genomic DNA. In order to identify potentially interesting cell lines, we used a "cobblestone area" (CA) assay (Ploemacher, et al., 1989) was initiated with BM taken from mice injected two days previously with 5-fluorouracil (5-FU). It has been suggested that CA colonies which appear after a prolonged culture period are derived from more primitive stem cells, possibly identical to some *in vivo* transplantable entities. (Ploemacher, et al., 1991). Therefore, a goal was in identifying cell lines which support such late arising CAs. Of 225 lines, 77 (34%) were capable of supporting limited *in vitro* hematopoiesis, while, consistent with the initial hypothesis, only 2% were able to maintain long-term (>6 weeks) hematopoietic CA activity. Subsequent studies with a selected subset of these lines, showed that the ability to effectively support *in vivo* reconstituting BM stem cells is infrequently observed. (Wineman, et al., 1996). Two out of sixteen cell lines maintained significant levels of long-term reconstituting stem cell activity for an *in vitro* culture period of three weeks. Several other cell lines supported low levels of such activity or transiently repopulating stem cells. The cell inoculum, in these studies, was whole BM which was not enriched for stem cell activity. Therefore, it was not possible to suggest that the effective stromal cell lines were directly supporting stem cell activity. This example demonstrates that a single clonal cell line, designated AFT024, can maintain quantitative levels of transplantable stem cell activity present in highly purified stem cell populations. These data were generated using a competitive repopulation assay system, which employs uncompromised competitor BM cells. The *in vitro*-maintained stem cells satisfy all criteria which currently define the most primitive stem cell population including the ability to reconstitute secondary recipients. This example also shows that the *in vitro* maintenance of primitive transplantable stem cells is compatible with the concurrent generation of large numbers of committed progenitors.

Methods

Mice. Timed-pregnant mice and 5-7 weeks old females (C57Bl/6J, Ly5.2) were purchased from the Jackson Laboratory (Bar Harbor, ME). Congenic C57Bl/6, Ly5.1 female mice were purchased from the National Cancer Institute (Frederick, MD). All mice were housed in the Princeton University Barrier Animal Facility, in autoclaved micro-isolator cages on ventilated cage racks. The animals received sterile, irradiated food, and acidified, autoclaved water ad libitum.

Stromal cell lines and culture conditions. The fetal liver stromal cell lines used in this study were derived as previously described. (Wineman, et al., 1996). Stromal cell lines were routinely cultured in Dulbecco's modified Eagles's medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 5×10^{-5} mol/L b-mercaptoethanol (2-ME), at 32°C, 5% CO₂, 100% humidity. Sera were obtained from Hyclone, Logan, UT. Other biochemical reagents were obtained from Sigma, St. Louis, MO. Two of the lines used in this study (2012 and 2018) were previously characterized for their ability to support long-term repopulating activity present in whole unfractionated BM. (Wineman, et al., 1996). The AFT024 cell line was identified as an additional long-term (>4 weeks) CA supporter. Subclones of 2012 and AFT024 were isolated and used in these studies. The AFT024 cell line has remained stable and demonstrated consistent stem cell supporting abilities for over 4 years.

Hematopoietic stem cell purification. Stem cells were purified from day 14 fetal livers essentially as described, (Jordan, et al., 1990) with the inclusion of c-kit expression as an additional parameter. Briefly, AA4.1⁺ cells were isolated by immunopanning on petri dishes coated with AA4.1 antibody (10 ug/mL). The AA4.1⁺ fraction has been shown to contain all repopulating stem cell activity present in day 14 fetal liver. (Jordan, et al., 1990) AA4.1⁺ cells were collected and stained with saturating concentrations of fluorescein isothiocyanate (FITC) labeled rat monoclonal antibodies to lineage markers (CD3, CD4, CD5, CD8, B220, Gr-1, Mac-1, and TER-119). The cells were simultaneously stained with phycoerythrin (PE) labeled Ly6A/E (Sca-1) antibody and biotinylated antibody to c-kit. The latter was developed with streptavidin allophycocyanin (APC). The AA4.1 hybridoma was a kind gift from Dr. J. McKearn, Monsanto, St. Louis, MO. AA4.1 antibody was purified by ImClone Systems Inc. New York, NY. The TER-119 antibody was initially obtained from Dr. T. Kina, Kyoto University, Japan and subsequently purchased from PharMingen, San Diego, CA. All

other antibodies were purchased from PharMingen. Streptavidin APC was purchased from Molecular Probes Inc., Eugene, OR. Three color fluorescence activated cell sorting for lineage negative to low ($\text{lin}^{-/\text{lo}}$), Sca-1⁺, c-kit⁺ cells was initially done on a dual laser Epics 753 cell sorter (Coulter Electronics, Hialeah, FA) interfaced with Cicero software
5 (Cytomation Inc., Fort Collins, CO) and subsequently on a multi-laser FACS Vantage with Cell Quest software (Becton Dickinson Immunocytometry Systems, San Jose, CA).

Enriched fractions of BM stem cells were obtained from C57Bl/6 Ly5.1 mice as described. (Okada, et al., 1992). Briefly, BM mononuclear cells were isolated by density centrifugation over Ficoll Hypaque (<1.077, Pharmacia, Piscataway, NJ). Lineage
10 negative or low staining cells ($\text{lin}^{-/\text{lo}}$) were obtained by magnetic bead depletion (anti-rat immunoglobulin coated beads, Dynal, Oslo, Norway) of mononuclear cells using the same lineage cocktail described herein above. The cells were further stained with antibodies to Sca-1 and c-kit as described herein above. Sorting for $\text{lin}^{-/\text{lo}}$, Sca-1⁺, and c-kit⁺ cells was accomplished with the Epics 753 as described above.

15 Stem cell/stromal cell cocultivation and cytokine-supplemented suspension culture. Stromal cell lines were seeded on tissue culture dishes that had been coated with 1% gelatin (Specialty Media, Lavallete, NJ) and were grown at 32⁰C, 5% CO₂, 100% humidity. Confluent monolayers were irradiated (20 Gy, ¹³⁷Cesium source, Gammacell 40, Nordion International Inc. Ontario, Canada) and cultured in modified Dexter (Dexter,
20 et al., 1984) media (DMEM, 10% FBS, 10% horse serum, 5X10⁻⁵ mol/L 2-ME, 1X10⁻⁷ mol/L hydrocortisone). For Dexter-LTC, enriched hematopoietic stem cells were added and the cultures were maintained at 37⁰ C, 5% CO₂, 100% humidity with weekly media changes. The specific numbers of purified stem cells added to stromal cell cocultures are given in the appropriate figure or table legends. In some experiments, week 4
25 AFT024/stem cell cultures were harvested and replated in limiting-dilution onto fresh, irradiated (20 Gy) AFT024 monolayers in 96-well trays (Dexter-LTC conditions). CAs were scored weekly (as described above), for an additional 5 weeks. Irradiated (20 Gy) 2018 monolayers in 96-well trays were used in limiting-dilution Whitlock-Witte assays (LD-WW) (Whitlock, et al., 1982) to assess stromal-dependent B-lymphopoiesis content
30 of both freshly purified and AFT024 cultured fetal liver stem cells. These cultures were established in RPMI media with 5% FBS, 2 mmol/L glutamine, 1 mmol/L Na pyruvate, and 5X10⁻⁵ mol/L 2-ME at 37⁰ C, 5% CO₂, 100% humidity. 2018 has been identified as a potent B-lymphopoiesis supporting line in W-W conditions. (Deryugina, et al., 1994).

Short-term cytokine-supported suspension cultures and short-term AFT024/stem cell cocultures were established in Iscove's Modified Dulbecco's Media (IMDM), 10% FBS, 1% BSA, 5×10^{-5} mol/L 2-ME at 37°C, 5% CO₂, 100% humidity. Cytokine concentrations: rmflk2/flt3-ligand (FL) 30 ng/mL, rmSteel factor (SL) 20 ng/mL, rhIL-6
5 10 ng/mL. FL was obtained from ImClone Systems Inc.; SL was purchased from Genzyme Corporation, Cambridge, MA. IL-6 was purchased from Upstate Biotechnology Inc., Lake Placid, NY.

Transplantation assays for hematopoietic stem cell activity. Competitive repopulation was used to measure stem cell activity present in both freshly isolated and cultured stem
10 cell populations. (Harrison, et al., 1993). This assay was performed using the congenic Ly5.1/5.2 mouse system (Morse, et al., 1987). Enriched stem cells were seeded onto irradiated stromal monolayers and maintained in Dexter-LTC conditions. At the end of 4-7 weeks, the cultures were harvested by vigorous trituration. Single cell suspensions were prepared by passage through 22-gauge needles, mixed with fresh congenic BM and
15 transplanted into lethally irradiated congenic mice (10 Gy, split dose 3 hours apart, 1 Gy/min, Gammacell 40). All purified fetal liver stem cells were from Ly5.2 mice. The competitor BM and recipients were Ly5.1. Purified BM cells were from Ly5.1 mice; in this experiment, Ly5.2 BM was used as competitor and Ly5.2 mice were used as recipients. In order to assess reconstitution, mice were periodically bled by capillary
20 puncture of the orbital venous plexus. Blood (0.1 mL) was collected into heparin-containing (10 U/mL) DMEM and the red blood cells were lysed with NH₄Cl (Mishell, et al., 1980). For the experiments described in Figure 15 and Table 4., the nucleated cells were divided into two fractions and stained with the appropriate biotinylated Ly5 antibody and developed with streptavidin-peridinin chlorophyll protein (PerCP) and; (1)
25 CD4-FITC, CD8-PE and (2) B220-FITC, Mac-1-PE, Gr-1-PE. Cells from each fraction were analyzed on an Epics Profile II, Coulter Electronics. For the experiments described in Figure 16 and Tables 2. and 3., nucleated cells were stained with directly conjugated lineage antibodies (CD4-PE, CD8-PE, Mac-1-FITC, Gr-1-FITC, and B220-APC) and biotinylated Ly5.2 antibody which was developed with streptavidin Texas Red (T.R.).
30 Four color analysis of stained cells was performed on either the Coulter Epics 753 or Becton Dickinson FACS Vantage with the appropriate software interfaces described above. Anti-Ly5.1 was a kind gift of Dr. H. Nakauchi, University of Tsukuba, Japan. Purified and biotinylated Ly5.2 antibody was originally obtained from a hybridoma

(AL14A2) kindly provided by Dr. G. Spangrude, University of Utah Medical Center, Salt Lake City, UT. In later experiments, the Ly5.2 antibody was purchased from PharMingen; CD4-PE, CD8-PE, and B220-APC were also purchased from PharMingen. Streptavidin-PerCP was purchased from Becton Dickinson Immunocytometry Systems. Streptavidin T.R. was purchased from Molecular Probes. Competitive repopulating units (CRU) per 10^5 were calculated according to the formula: (Harrison, et al., 1990).

$$\text{CRU}/10^5 = \frac{\% \text{Ly5 positive cells}}{100 - \% \text{Ly5}} \times \frac{\text{cell number competitor BM}}{\text{number of test cells}}$$

10

Reconstitution values of less than 2% of the test Ly5 donor allele were not considered sufficiently above background for calculation of CRU.

Retransplantation potential of stem cells was assessed by secondary transplantation. Mice from the experiment presented in Figure 14. were sacrificed 60 weeks after transplantation, BM was harvested and stained with antibody to Ly5.2 followed by streptavidin T.R. Ly5.2⁺ cells were collected by cell sorting (Coulter Epics 753) and used to transplant lethally irradiated secondary Ly5.1 recipients. Marrow from mice in the Control and AFT024 groups was used to transplant mice in both radioprotection and competitive repopulation assays. Ly5.2⁺ BM from the primary 2012 transplants was used only in radioprotection assays. Primary 2018 mice did not contain sufficient Ly5.2⁺ cells for secondary transplantation. Transplanted mice were bled and analyzed by 4-color flow cytometry for the presence of Ly5.2⁺ cells and multilineage reconstitution as described above.

In vitro hematopoietic progenitor cell assays. The progenitor content of freshly purified hematopoietic stem cell populations and AFT024/stem cell cocultures was assessed using a variety of *in vitro* assays. All of the following assays were accomplished with fetal liver stem cells enriched as described herein above. To determine the time course of CA development, enriched stem cells were seeded onto irradiated AFT024 monolayers (300-600 cells/well in 12-well trays). CA development was followed over time and characteristic clusters were quantitated as described above. At different time points of the stem cell/AFT024 cultures, individual wells were harvested and replated into cytokine-supplemented semisolid clonogenic progenitor assays (CFU-C). The cytokine-enriched (rmIL-3 10 ng/mL, rhIL-6 10 ng/mL, rmSL 50 ng/mL, Epo 3 U/mL) methyl-

30

cellulose mixture was purchased from Stem Cell Technologies Inc., Vancouver BC, Canada. Colonies were scored after 8-14 days of culture at 37⁰, 5% CO₂, 100% humidity according to established criteria. (Testa, et al., 1993). Colonies that reached >1mm in size after 8 days and which contained erythroid bursts and multiple myeloid cell lineages including megakaryocytes were scored as high-proliferative potential-mixed lineage colonies (CFU-HPP-Mix). (Lowry, et al., 1995). Lineage content of typical colonies was determined by Wright's/Giemsa staining of cytopsin slide preparation from individual colonies. Colony assays were also done with 10³ freshly purified cells. The CFU progenitor contents of the AFT024 cocultures were normalized to an initial input of 10³ stem cells. To assess the ability of the AFT024 cell line to maintain primitive lymphoid progenitors, 4 week cocultures were plated into LD-WW assay on 2018 cells as described above. Resulting pro-B cell colonies were scored after 7 days. The cell number in individual wells (96-well trays, 8 wells/cell number) was normalized from the original number of purified stem cells that initiated the coculture, i.e. stem cell equivalents/well. As calculated from the line of best fit, the cell number at 37% negative wells is the frequency of pro-B cell colony initiating cells in the starting population. (Taswell, 1981). In a similar manner, the frequency of CA initiating cells in week 4 stem cell/AFT024 cocultures was also determined by replating them in limiting-dilution onto fresh, irradiated AFT024 monolayers in 96-well trays. CA were scored as described above, at 1, 2, 3, 4, and 5 weeks after replating in Dexter-LTC. The resulting frequencies were calculated as described herein above for the LD-WW assays and are also expressed in relationship to the number of stem cells that seeded the initial cultures (stem cell equivalents).

Results

In vivo and in vitro assays for stem cell activity maintained by stromal cell lines. Highly enriched stem cell populations were used to initiate cultures supported by single stromal cell lines. The focus was on one cell line, AFT024, which exhibited particularly potent stem/progenitor cell support. Two other stromal cell lines, 2012 and 2018, (Wineman, et al., 1996) were included in some experiments. In order to more rigorously establish the clonality of these lines, they were subcloned by limiting-dilution. All subclones obtained from a given cell line contained the same proviral integrant position as the parental cell line. The AFT024 cell line was evaluated both for its ability to maintain *in vivo* competitive repopulating stem cells, as well as a broad spectrum of

stem/progenitor cells defined by a variety of *in vitro* assays. The *in vivo* assays focused on the ability of stem cells, cultured for extended time periods (4-7 weeks), to permanently reconstitute multi-lineage hematopoiesis in transplanted hosts. The *in vitro* assays included the enumeration of CA appearing over time in the initial cultures as well as the quantitation of stem/progenitor cells which can form colonies in cytokine-supplemented replating assays. Cells from four-week AFT024/stem cell cocultures were also assayed by limiting-dilution for the content of progenitors capable of initiating secondary CAs on AFT024 or B-lymphopoiesis in Whitlock-Witte cultures supported by 2018.

AFT024 maintains quantitative levels of long-term *in vivo* repopulating stem cell activity. One line of investigation inquired if and at what levels *in vivo* transplantable stem cell activity was present in four to seven week-old cultures initiated with highly enriched stem cells and supported by AFT024, 2012 or 2018. Purified day 14 fetal liver cells (AA4.1⁺, lin^{-lo}, Sca-1⁺, c-kit⁺) and adult BM cells (lin^{-lo}, Sca-1⁺, c-kit⁺) were used as sources of stem cell activity. Both of these populations are about 1000 to 1500-fold enriched for stem cell activity, as measured by competitive repopulation. (Harrison, et al., 1993). The Ly5.1/Ly5.2 congenic system was utilized for all competitive repopulation studies. (Morse, et al., 1987). The data presented in Figure 14 demonstrate that the cultures supported by AFT024 contain stem cell activity at levels quantitatively identical to those present in the uncultured purified populations. In this experiment, individual Ly5.1 mice received 10³ freshly purified Ly5.2 cells or the cultured equivalent of 10³ purified Ly5.2 cells. Each mouse also received 10⁶ Ly5.1 competitor BM cells. The percentage of Ly5.2 positive peripheral blood cells was approximately equal in both groups of recipient animals. Moreover, the cultured stem cell activity is as effective as freshly purified activity for *in vivo* periods of greater than one year. The data in Figure 14 also show that the 2018 cell line is completely ineffective in maintaining highly purified stem cell activity while the 2012 cell line supports intermediate levels of repopulating activity. The data presented in Table 4A. provide quantitative competitive repopulating unit (CRU) value calculations as well as the results of multiparameter lineage analyses. The extremely low levels of reconstitution by 2018-cultured stem cells precluded lineage analysis. The CRU values of the AFT024-cultured and freshly purified populations are nearly identical. Moreover, both fresh and AFT024-cultured stem cells reconstitute myeloid and lymphoid cell populations to a similar degree. In order to further access the

supporting activities of AFT024 and 2012 we utilized a 10-fold lower number of fetal liver stem cells from two separate purifications to initiate the cocultures. The cultures were continued for 4-7 weeks, harvested, and used in competitive repopulation studies. Each recipient received the cultured equivalent of 100 purified Ly5.2 stem cells plus
5 4X10⁵ Ly5.1 competitor BM cells. A total of twelve mice were transplanted with AFT024 cocultures (four each after four, five and seven weeks of LTC). The parental AFT024 line was used in the 4 week group and two different subclones were used to support the five and seven week cultures. The AFT024-cultured Ly5.2 stem cells contributed to 20-30% of peripheral blood cells in these recipients while cells cultured on
10 2012 demonstrated more limited *in vivo* function (Table 4B). The 2012 cultures were done with two subclones of the parental line and were maintained for four weeks prior to harvest and transplant (four mice/subclone). The data utilizing different cultures time or subclones did not vary significantly from each other and are presented together in Table 4B. An additional experiment was undertaken using enriched BM which was cultured on
15 AFT024 and 2018 for six weeks. In this experiment, BM was purified from Ly5.1 congenic mice. Each Ly5.2 recipient mouse in this study received 100 freshly purified cells, or the cultured equivalent of 100 purified cells. Both groups received 10⁵ Ly5.1 competitor BM cells per mouse. Data analysis, for the presence of Ly5.1⁺ cells at 4 months after transplant, is presented in Table 4C. For 6 weeks of culture, AFT024 cells
20 maintained quantitative levels of reconstituting activity present in 100 purified BM stem cells. The 2018 cell line failed to maintain stem cell activity.

These studies were extended to include secondary transplantation as an additional assay for primitive stem cells. BM cells were harvested from the primary recipients of fresh and cultured fetal liver stem cells (see Figure 14. and Table 4A.) and the Ly5.2
25 positive, fetal liver-derived fraction was collected by cell-sorting. Secondary radioprotection and competitive repopulation transplants were performed. The data are presented in Table 5. The secondary recipient repopulating activities are nearly identical for the AFT024 cultured stem cells and the non-cultured controls. Lineage analysis of the Ly5.2 cells in the secondary recipients revealed similar numbers of myeloid and
30 lymphoid cells derived from both AFT024-cultured and non-cultured stem cells. Some level of secondary reconstituting cell activity was observed.

Additional experiments were also performed to determine where the levels of stem cell activity present in long-term AFT024 cocultures were compared to those present

in short-term cytokine stimulated cultures or in short-term AFT024-supported cultures (Figure 15). Purified fetal liver cells were seeded onto an AFT024 monolayer and maintained in Dexter-LTC conditions for 5 weeks. Simultaneously, the same numbers of purified cells were cultured for 5 days with; (1) different cytokine combinations or (2) on AFT024. The transplantable activity in the cultured cells was then assayed by competitive repopulation. Each mouse received the cultured equivalent of 600 stem cells together with 4×10^5 Ly5.1 competitor BM cells. It is evident from the data that the levels of *in vivo* repopulating activity present in the long-term AFT024-supported cultures are much greater than those remaining after a short-term cytokine-supported culture period. Of interest also is that short-term AFT024 stem cell cocultures do not maintain significant *in vivo* reconstituting activity. In fact, these levels of stem cell activity are identical to the levels seen in the cytokine-supported cultures.

In vitro stem/progenitor populations are expanded by AFT024. AFT024/stem cell cocultures have vigorous hematopoiesis throughout the entire *in vitro* culture period. This is reflected in the large numbers of relatively mature hematopoietic cells which are produced throughout the culture period. In addition, CA colonies are observed throughout the culture period. Figure 16 shows a time course of CA appearance with purified fetal liver stem cells (3 separate experiments). After 28 days in culture, approximately one in every twenty input stem cells is capable of proliferating into a CA. In addition, CA appearance over time follows a biphasic distribution, with many CA observed early in the culture period. In order to enumerate the various classes of stem/progenitor cells present in AFT024 cocultures, we performed a series of *in vitro* replating experiments. These included the quantitation of: (1.) progenitor cells capable of colony-formation in cytokine-supplemented semisolid cultures (CFU assay), (2.) progenitor cells capable of initiating secondary CA in limiting-dilution AFT024 cultures and (3.) progenitor cells which can initiate B-lymphopoiesis in LD-WW cultures. In all of these experiments the primary cultures were initiated with purified fetal liver cells. All data presented below are normalized to an initial input of 10^3 purified cells (CFU assay) or the actual number of initial input stem cells (limiting-dilution assays).

Shown in Figure 17 are the numbers and types of cytokine responsive CFU progenitors present at various times in the AFT024-supported cocultures. Production of CFU is evident at all time points. However, the content is especially high after four weeks, representing a 5-7 fold increase/expansion when compared to the content in the

freshly purified populations. The content of more primitive progenitors (CFU-HPP-Mix) is increased by 12-fold. These HPP-Mix colonies often reach a size of 2 mm in 8 days and contain large numbers of erythroid bursts and megakaryocytes. Interestingly, there does not appear to be a correlation between CA number and CFU content at different culture times. This is most apparent at day 6, when CA numbers are at their peak but the progenitor content is similar to that observed in non-cultured stem cells. Furthermore, there is no correlation between CFU content and the absolute numbers of maturing hematopoietic cells present in a given culture.

Next, the content of primitive B-lymphoid progenitors present in the AFT024/stem cell cultures was determined. This was accomplished by plating cells from the four-week AFT024 cocultures into LD-WW assays over the 2018 stromal cell line. Two experiments with freshly purified stem cells and AFT024-cultured stem cells showed that the frequency of pro-B cell progenitors is expanded 10-fold in AFT024 cultures compared to the frequency observed in the freshly purified input population (day 0 frequency 1 in 11.0, $r^2 = 0.98$; day 28 AFT024-cultured frequency 1 in 1.1, $r^2 = 0.97$).

In order to measure the content or frequency of progenitor cells capable of initiating secondary CA, 4 separate, four-week AFT024/stem cell cocultures were replated in limiting-dilution onto fresh AFT024 monolayers. CAs were scored after one week. The data are presented in Figure 18A. Large numbers of secondary CAs were observed. When normalized to the stem cell numbers used to initiate the primary cultures (stem cell equivalents, see Methods), the frequency of these progenitors is 1 in 3 to 4. Figure 18B. shows data from one of the 4 experiments presented in Figure 18A., where the quantitation of secondary CAs was extended for 4 more weeks. The frequency of CA decreases slowly over time (1 in 19 after an additional 4 weeks), approximating the frequency seen in the primary cultures at four weeks. In summary, our *in vitro* replating assays collectively demonstrate a significant expansion of primitive progenitor populations in 4 week AFT024 cultures. In these same cultures there is no decrease in the levels of transplantable stem cell activity present in the total hematopoietic cell population.

Discussion

In this example, it was demonstrated, that the AFT024 stromal cell line can maintain quantitative levels of *in vivo* repopulating stem cells for at least 7 weeks of *in vitro* culture. Highly enriched stem cell populations in low numbers (100 cells) were

used and cell activity was measured in a stringent, competitive repopulation assay system. The cultured stem cell activity satisfies all *in vivo* criteria normally ascribed to the most primitive stem cell compartment; (1) long-term engraftment ability, (2) multilineage potential and (3) the ability to repopulate secondary recipients. In addition, the studies with the low culture initiating stem cell numbers for both BM and fetal liver, imply that AFT024 stromal cells exert their supportive effects in a direct manner. These studies represent a clear example of an *in vitro* system capable of directly supporting the most primitive stem cell compartment.

Other data revealed that stromal cell lines isolated from a single tissue source are heterogeneous with respect to their abilities in maintaining long-term repopulating stem cells. (Wineman, et al., 1996). It was speculated that the rare cell lines which were effective in supporting *in vivo* reconstituting stem cells may represent immortalized components of *in vivo* stem cell niches. However, the data argue for the necessity of using purified stem cell populations in order to support such a hypothesis. Specifically, the 2018 cell line maintained transiently reconstituting activity present in unfractionated BM. However, 2018 fails to maintain measurable repopulating activity when cultured for 4-6 weeks with highly purified BM or fetal liver stem cells. Similarly, in experiments using purified cells, there was a failure to show robust levels of reconstituting stem cell activity in cultures supported by CFC034, the most effective cell line in the whole BM studies (Wineman, et al., 1996). The 2012 cell line which was reported (Wineman, et al., 1996) to be effective in maintaining long-term repopulating whole BM derived stem cells is only partially effective in the present studies. Moreover, only some subclones of 2012 display such activity (in spite of identical proviral integration positions in all subclones). Recently, studies have shown that the S17 cell line which consistently supports the stem cell activity present in whole BM, (Wineman, et al., 1993) is not similarly effective in the maintenance of purified BM stem cells. (Szilvassy, et al., 1996). Taken together with these current data, the previously observed stem cell supporting stromal cell activities may reflect the action(s) of indirect mechanisms and therefore do not permit the identification of cellular stem cell niche components. One previous study has shown that the Sys1 stromal cell line can maintain high levels of transplantable activity present in purified BM. (Szilvassy, et al., 1996). The competitor cells in that study were compromised by prior serial transplantation. Moreover, the culture period was extended for only two weeks and effective maintenance required the addition of exogenous

leukemia inhibitory factor. In contrast, AFT024 is a cell line that provides a direct-acting long-term stem cell supporting environment without the addition of exogenous factors.

In addition to recovering net input levels of transplantable activity from AFT024 supported cultures, significantly expanded populations of primitive progenitor cells were also obtained. CFU-HPP-Mix progenitors are expanded by 12-fold after 4 weeks of culture and the numbers of stromal-dependent pro-B lymphoid progenitors are similarly amplified. This suggests that the AFT024-mediated process of stem cell maintenance is in reality a dynamic phenomenon. Specifically, during the first portion of the culture period, the majority of transplantable stem cell activity may be lost, through differentiation or cell death. The remaining primitive stem cells may expand to yield input levels of transplantable activity as well as increases in the numbers of more committed progenitors. One hypothesis is that short-term AFT024 supported cultures should contain reduced levels of transplantable stem cell activity. Figure 17 support this hypothesis. The standard Dexter-type media used in parallel long-term AFT024 cocultures was not used in these short-term cultures. However, in another short-term experiment, utilizing Dexter-LTC media, a similarly dramatic reduction in stem cell activity was observed after 4 days of culture on AFT024.⁶³ These observations are intriguing because they suggest that the AFT024 cell line is able to facilitate some degree of *ex vivo* transplantable stem cell proliferation and expansion. Indeed, in other studies it was shown that AFT024 can support colony formation initiated by single purified stem cells with B and T-lymphoid, myeloid and erythroid potentials. Moreover, the data suggest that is possible to efficiently introduce retroviral markers into transplantable stem cells at various times during AFT024 cocultures. Extension of such marking experiments and an analysis of proviral integration patterns will be necessary to rigorously ascertain if self-renewal replication is occurring during these coculture periods.

The ability of AFT024 to maintain the most primitive stem cell compartment while generating and expanding at least some less primitive members of the stem/progenitor cell hierarchy raises interesting issues regarding the nature of stem cell niches. The present invention suggests that microenvironmental niche models which postulate distinct cellular entities responsible for stem cell self-renewal and other cellular entities which support the generation of committed progenitor cells may be overly simplified. (Uchida, et al., 1993). Quite clearly a single microenvironmental cell type represented by AFT024 is sufficient for keeping stem cells in an undifferentiated state as

well as allowing commitment and progenitor expansion to take place. A hallmark feature of a stem cell niche is the ability to facilitate the generation of the entire stem/progenitor cell hierarchy from very primitive cells. Therefore, the main functional role of such niches may be to provide an environment which permits the production of the correct numerical balance of more and less primitive stem/progenitor cell entities. This model contains several testable hypotheses. The most important is that *in vitro* stem cell maintenance should not be interpreted, literally, as the maintenance of quiescent cells but rather as a phenomenon which results from a balance of self-renewal and commitment decisions which occur during stem cell division.

A cytokine cocktail of IL-6, SL, and FL is not effective in maintaining fetal liver stem cell activity. It was shown that RNA transcripts for these and 10 other cytokines are present in AFT024, but they are also detected at similar levels in non-supporting lines such as 2018. (Wineman, et al., 1996). These observations suggest the existence of novel AFT024-derived molecules which may act on stem cells. Indeed, using a subtractive hybridization molecular cloning strategy, a number of candidate molecules have been identified. Two of these molecules contain EGF-like repeat motifs that are most closely related to those found in the Notch/Notch-ligand family. Interestingly, one of these molecules appears to have activity on primitive stem cell populations. (Moore, et al., 1997). (See, Examples presented herein above.

Table 4. Multilineage stem cell activity, high CRU levels maintained in AFT024 cultures
A. With 1000 purified fetal liver cells

Cells	Week	Lineage Contribution, % Ly 5.2 Cells				Total % Ly 5.2	CRU/10 ⁵
		CD4	CD8	B220	Myeloid		
Control	5	3.3±1.5	4.4±1.7	48±5.1	32±6.0	28±1.3	386±24
	12	27±10	24±8.7	63±8.9	43±9.2	44±3.2	800±102
	24	42±6.9	34±4.1	71±4.4	62±1.4	48±4.8	984±171
AFT024	5	7.7±6.2	6.0±3.5	58±6.6	36±4.7	39±2.1	641±54
	12	44±7.1	38±6.4	70±5.6	53±3.6	48±1.4	942±51
	24	56±3.8	49±3.1	72±2.4	64±3.7	56±1.9	1270±96
2012	5	0.8±0.5	1.4±0.3	6.8±3.6	14±11	7.4±2.3	52±8.5
	12	9.1±8.4	10±6.9	14±9.3	23±20	14±4.5	194±68
	24	17±3.8	18±2.0	23±5.9	32±6.4	21±4.4	196±71

B. With 100 purified fetal liver cells

Cells	Week	Lineage Contribution, % Ly5.2 Cells			Total	
		Granulocytes	B-cells	T-Cells	% Ly5.2	CRU/10 ⁵
AFT024	15	33±4.6	41±6.0	25±5.3	28±4.6	1888±315
	26	17±4.4	18±4.8	18±5.9	24±5.9	1825±571
	46	22±5.7	19±6.6	15±6.7	20±6.5	1595±628
2012	15	32±7.8	28±7.2	20±5.2	24±6.6	1390±382
	26	10±2.8	9.7±3.2	9.7±3.2	14±3.9	699±176
	46	11±2.3	6.7±1.6	8.9±3.2	9.0±2.0	409± 80

C. With 100 purified bone marrow cells

Cells	No. of mice	Lineage Contribution, % Ly 5.1 Cells				Total	
		CD4	CD8	B220	Myeloid	% Ly 5.1	CRU/10 ⁵
Control	8	38±7.3	30±5.8	43±6.7	31±4.8	32±6.5	583±172
AFT024	7	20±9.8	18±8.8	20±8.7	24±11	22±10	490±296

- 5 A. AA4.1⁺ day 14 fetal liver cells (Ly 5.2) were further purified for a lin^{-lo}, Sca-1⁺, ckit⁺ stem cell surface phenotype. Fresh purified control cells, 10³, were transplanted with 10⁶ Ly5.1 competitor marrow (n=6 mice). From the same purification, 10⁴ cells were cocultured with stromal cell lines for 4 weeks. Subsequently, 10% of each culture was transplanted per mouse (n=8 mice/stroma) together with 10⁶ competitor BM cells. The contribution to each lineage in peripheral blood is expressed as the percent of the total specific lineage population that was Ly5.2⁺. CRU/10⁵, relative enrichment of competitive
- 10 repopulating units. Data are presented ±SEM.

- B. The lineage and CRU content of low numbers of enriched fetal liver stem cells maintained on AFT024 and 2012 were determined. 500 stem cells were maintained in Dexter-LTC for 4-7 weeks over irradiated monolayers. 20% of each culture was used to transplant groups of 4 mice (i.e. each mouse received the
- 15 equivalent of 100 stem cells that initially seeded the cultures) combined with 4X10⁵ competitor BM cells. Data are presented as ±SEM.

- C. Ly 5.1 BM cells with a lin^{-lo}, Sca-1⁺, and c-kit⁺ cell surface phenotype were purified. 100 fresh cells per mouse (Control) were transplanted with 10⁵ Ly 5.2 competitor BM cells. One thousand of the same
- 20 purified cells were cocultured with stromal cell lines for 6 weeks. The cultures were then harvested and 10% of each culture was transplanted per mouse with competitor. Data are from peripheral blood samples taken 4 months after transplant and are presented ±SEM.

Table 5. LTRSC maintained on AFT024 are able to repopulate secondary recipients at levels comparable to non-cultured stem cells.

Group	Weeks	% Ly 5.2 ⁺ peripheral blood cells	
		Radioprotection	Comp. Repopulation
Control	6	21 (1)	2.6±0.6 (4)
	22	54 (1)	5.6±1.6 (4)
AFT024	6	13±0.2 (4)	1.3±0.9 (8)
	22	44±15.1 (3)	4.3±0.9 (8)
2012	6	6.7±2.9 (4)	ND
	22	14±2.2 (4)	ND

The retransplantation potential of LTRSC in primary recipients of stromal cell cultured stem cells was studied in secondary recipients. 60 weeks after transplant, primary mice (see Figure 14, Table 4A.) were sacrificed, BM harvested and stained with antibody to Ly5.2. Ly5.2⁺ cells were collected by cell sorting and used to transplant secondary recipients (congenic Ly5.1 mice). Control and AFT024 groups were transplanted with 1.5×10^6 Ly5.2 cells/mouse for radioprotection (4 mice/group were transplanted) and 7.5×10^5 Ly5.2 cells + 7.5×10^5 Ly5.1 cells for competitive repopulation (4 mice for the Control group and 8 mice for the AFT024 group). 2012 mice were transplanted with 3×10^5 Ly5.2 cells/mouse (4 mice). Weeks are the times after transplant that the mice were analyzed. (n), number of mice surviving/group; ND, not done. Data are presented ±SEM.

EXAMPLE 4: Hematopoietic activity of a stromal cell transmembrane protein containing epidermal growth factor-like repeat motifs

Primitive hematopoietic stem cells are closely associated with discrete *in vivo* microenvironments. These "niches" are thought to provide the molecular signals that mediate stem cell differentiation and self renewal. The fetal liver microenvironment was dissected into distinct cellular components by establishing an extensive panel of stromal cell lines. One particular cell line maintains repopulating stem cells for prolonged *in vitro* culture periods. A subtraction cloning strategy has yielded a cDNA which encodes a cell surface glycoprotein with a restricted pattern of expression among stromal cell lines. This

molecule, previously identified as delta-like/pre-adipocyte factor-1, contains epidermal growth factor-like repeats which are related to those in the notch/delta/serrate family of proteins. The potential role of this molecule in hematopoietic stem/progenitor cell regulation was investigated. The delta-like protein displays activity on purified stem cells by promoting the formation of "cobblestone areas" of proliferation. These cobblestone areas contain both primitive high-proliferative potential progenitors and *in vivo* repopulating stem cells.

Introduction

The positive and negative regulatory mechanisms that govern the proliferation, self-renewal and differentiation of primitive hematopoietic stem cells are complex and poorly understood (Ogawa, M. (1993) *Blood* 81, 2844-2853). Numerous cytokines have been identified that, when used *in vitro* appear to act directly on purified stem cells by promoting proliferation and differentiation. However, attempts to demonstrate the *in vitro* maintenance and/or expansion of transplantable pluripotent stem cells using defined cytokine combinations have been largely unsuccessful (Knobel, K. M., et al., 1994; Peters, S. O., et al., 1995). Moreover, it is unclear to what extent any currently identified cytokines reflect mechanisms that are responsible for regulating normal, *in vivo*, stem cell behaviors. It is widely accepted that *in vivo*, stem cells are intimately associated with discrete microenvironmental "niches" (Wolf, N. S. 1979). Such niches are likely sources for the molecular signals which collectively mediate the differentiation and self-renewal of stem cells. Indeed, it has long been possible to demonstrate that preestablished stromal cell monolayers derived from hematopoietic tissues can support long-term hematopoiesis *in vitro* (Dexter, T. M., Allen, T. D. & Lajtha, L. G. 1977). The long-term nature of these cultures, together with the continuous production of committed progenitor cells suggest that both self-renewal and commitment decisions can occur *in vitro*. At the cellular level, the hematopoietic microenvironment consists of numerous distinct cell types. Previous studies have shown that this cellular heterogeneity reflects a similarly broad heterogeneity in terms of hematopoietic supportive abilities (Deryugina, E. I., et al., 1994). Some cloned stromal cell lines can support stem cell activity *in vitro*, while others are ineffective. Similarly, distinct stromal cell types appear to influence the outcomes of stem cell differentiation processes (Friedrich, C., et al., 1996). Recent studies have shown that stromal cell lines that efficiently maintain long-term transplantable stem cells *in vitro* for prolonged intervals represent a small fraction of the total stromal cell population (Wineman, J., et al., 1996). A fetal liver stromal cell line, AFT024 was identified which maintains high levels of

transplantable multilineage stem cell activity for extended *in vitro* culture periods (Moore, K. A., Ema, H. & Lemischka, I. R. (1997) (See, Example 3). The stem cells used to initiate these cultures are highly purified. It is, therefore, likely that in this system the mechanisms that mediate stem cell maintenance do so by acting directly on primitive stem cells. A different fetal liver-derived cell line, 2018, fails to maintain long-term repopulating stem cell activity. A PCR-based RNA expression analysis of 13 cytokines reveals qualitatively identical expression patterns in AFT024 and 2018. It was therefore, hypothesized that the hematopoietic supportive ability of AFT024 is, at least in part, mediated by novel gene products not expressed in 2018.

Using a subtractive hybridization approach a number of cDNA clones were identified which were specifically expressed in AFT024 but not in 2018. The sequence of one cDNA was identical to a molecule that encodes a transmembrane protein that contains six epidermal growth factor (EGF) repeat motifs. The EGF-like repeat sequences of this molecule, variously known as delta-like (dlk) (Laborda, J., et al., 1993), preadipocyte factor-1 (Smas, C. M. & Sul, H. S., 1993), and stromal cell protein-1 (Genbank, D16847), are most closely related to those present in the notch/delta/serrate family of signaling proteins. In *Drosophila* and *Caenorhabditis*, these molecules are required for correct cell-fate specification decisions in a variety of tissues (Rebay, I., et al., 1991). Vertebrate homologs of the notch/delta/serrate family have been identified (Ellisen, L. W., et al., 1991; Bettenhausen, B., et al., 1995; Lindsell, C. E., et al., 1995). While the exact functional relationship of dlk to the activities of this family of molecules is unclear, in one *in vitro* study, it has been shown to block adipocyte differentiation (Smas, C. M. & Sul, H. S., 1993). No studies have been reported that demonstrate a hematopoietic function for dlk. Expression analyses and these observations show a limited temporal pattern of dlk expression during murine fetal development which coincides with the time period of hematopoietic stem cell expansion (Smas, C. M. & Sul, H. S., 1993).

Functional studies were undertaken to determine if dlk can act as a hematopoietic regulator. This molecule affects highly enriched stem cell populations by promoting "cobblestone area" (CSA) colony formation in dexter-type stromal cocultures. These CSA colonies contain an expanded population of primitive, high proliferative potential myeloid-erythroid progenitors. These cultures also contain stem cells capable of *in vivo* engraftment at levels equivalent to those present in parallel AFT024 supported cultures. It is proposed that dlk represents one molecular component responsible for the hematopoietic supportive

ability of AFT024. As such, dlk may define a novel molecular pathway of stem cell regulation by the microenvironment.

Materials and Methods

Cell Lines and Culture. The fetal liver stromal cell lines used in this study were
5 derived as previously described (Wineman, J., et al., 1996). Cells were routinely cultured in
DMEM containing 10% fetal bovine serum (FBS) and 50 μ M b-mercaptoethanol (2-ME),
and maintained at 31⁰-33⁰C, 5% CO₂. For long-term cocultures with hematopoietic stem
cells, confluent monolayers were irradiated (20 Gy), placed in modified Dexter media
(DMEM, 10% FBS, 10% horse serum, 50 μ M 2-ME, 0.1 μ M hydrocortisone) and
10 maintained at 37⁰ C, 5% CO₂ with weekly media changes. NIH3T3 cells were obtained from
ATCC.

dlk Expression Analysis. Total RNAs from stromal cell lines were poly A⁺ selected,
Northern blotted, and hybridized to ³²P-labeled probes according to standard protocols
(Sambrook, J., Fritsch, E. F. & Maniatis, T, 1989). A 600 bp dlk-cDNA clone from the
15 AFT024 subtracted library was used as a probe. cDNA templates for RT-PCR were
prepared according to manufacturers' protocols (GIBCO/BRL). Oligonucleotide primers
were: sense 5'- GACCCAGGCTGCCCC-3' (SEQ.ID.No.:85) and antisense 5'-
GGTACTCTTGTTGAG-3' (SEQ.ID.No.:86). For analysis of dlk expression at the protein
level, antisera specific for dlk was generated by immunizing rabbits with a Flag-dlk fusion
20 protein (described below). Resultant antibodies were purified by affinity chromatography.
Cell surface expression of dlk in stromal cell lines was accomplished by flow cytometry.
Cells were incubated with dlk antibody and a similarly prepared irrelevant control antibody.
Specific labeling was developed by donkey anti-rabbit-fluorescein isothiocyanate (Jackson
ImmunoResearch). Stained cells were analyzed on a Becton Dickinson FACScan using Cell
25 Quest software.

dlk Fusion Protein Preparation. The expression plasmid pCD4-Ig contains cDNA for
the extracellular domain of human CD4 fused to genomic sequences of the human
immunoglobulin heavy chain (Zettlmeissl, G., et al., (1990). cDNA for CH2-CH3 of human
IgG₁ (Goodwin, R. G., et al., 1990) was cloned into *Eco*RI and *Not*I sites of pcDNA3
30 (Invitrogen) to give the plasmid KB52.3.2. cDNA encoding the extracellular domain of dlk
was obtained by RT-PCR with primers BP 151 and BP 152 using total RNA from NIH 3T3
cells as template. The resulting PCR fragment was cloned into KB52.3.2 via *Hind*III and

*Eco*RI sites to obtain the soluble dlk-Ig expression plasmid. pdlk-Ig or pCD4-Ig were transfected into NIH3T3 cells together with pSVNeo and stable clones were isolated. Soluble CD4-Ig and dlk-Ig fusion proteins were harvested and then purified by affinity chromatography on HiTrap Protein G-sepharose (Pharmacia). Primers: sense BP 151,
 5 5'GAGGGTACCAAGCTTCGTGGTCCGCAACCAGAAG-3' (SEQ.ID.No.: 87);
 anti-sense BP 152, 5'-CTCAGATCTGAATTCGGCCTGTCCCTCGGTGAGGAG-3'
 (SEQ.ID.No.: 88).

Flag-dlk fusion protein was used to immunize rabbits for the production of dlk antiserum. The protein expression plasmid pcDNA3-Flag is a modification of the plasmid
 10 pcDNA3 (Invitrogen) and contains the coding region for the Flag peptide (DYKDDDDKI) (Hopp, T. P., et al., 1988) as well as a *Bgl*III restriction site. A cDNA fragment encoding the extracellular domain of dlk was obtained by RT-PCR using RNA from NIH3T3 cells. Primers: sense BP 155,
 5'-GACAAGATCTCAGCTGAATAGCGACCCACCCTGTG-3' (SEQ.ID.No.: 89);
 15 antisense BP 154,
 5'-GCATCTAGAGCGGCCGCTCAGGCCTGTCCCTCGGTGAGGAG-3' (SEQ.ID.No.: 90). The PCR fragment was ligated into pcDNA3-Flag to yield pFlag-dlk. pFlag-dlk was transfected into cos cells. Purification of the Flag-dlk protein from cos-conditioned media was performed according to manufacturer's directions using the Flag monoclonal antibody,
 20 M1, immobilized on agarose (International Biotechnologies).

Plasmid Constructs and Stable Transfection. Full-length murine dlk cDNA was obtained by RT-PCR with primers BP 151 (see above) and antisense BP 200:
 5'GCATCTAGAGCGGCCGCGAACGCTGCTTAGATCTCCT-3' (SEQ.ID.No.:91), using
 total RNA from NIH3T3 cells as template. The product was subcloned into the vector
 25 pCRII (Invitrogen) and then cloned into a retroviral expression vector (Kitamura, T., et al., 1995), (G. Nolan, Stanford University), via the primer-encoded *Hind*III and *Not*I sites. Supercoiled plasmid was transfected into BFC012 stromal cells together with the pZeo (Invitrogen) selectable marker and selected in 50 ug/ml Zeocin (Invitrogen). BFC012 cells
 also were transfected with pZeo alone and selected as above. Clones from both selected
 30 populations were isolated and all remaining colonies (100-200 per dish) were pooled and expanded as populations.

Hematopoietic Stem Cells and *In Vitro* Hematopoietic Assays. Hematopoietic stem cell populations were derived from wild type, Ly5.2-C57Bl/6J (Jackson Laboratories), day

14 fetal liver, enriched for the AA4.1⁺, Sca-1⁺, *c-kit*⁺, and *lin*^{lo/-} phenotype, by immunopanning and fluorescence-activated cell sorting as described (Jordan, C. T., et al., 1995). Adult bone marrow (BM) was used directly after density centrifugation and immunomagnetic bead depletion or was further enriched for Sca-1⁺, *c-kit*⁺, *lin*^{lo/-} cells by
5 flow cytometry as described (Okada, S., et al., 1992). Cell sorting and data analysis was accomplished with a Becton Dickinson FACS Vantage using Cell Quest software. Stromal cell/stem cell cocultures were initiated in 12-well trays with 300-1,000 enriched stem cells per well. Cobblestone areas were quantitated by inverted-phase microscopy as described (Ploemacher, R. E., et al., 1991). Clonogenic progenitor assays were performed with either
10 freshly purified stem cells or cells harvested from the stromal cocultures. These were cultured in cytokine-containing semisolid media according to the manufacturer's recommendations (Stem Cell Technologies, Vancouver, BC). Soluble dlk and control fusion proteins were added to semisolid progenitor assays at concentrations of 0.1, 0.5 and 1.0 ug/ml and also to BFC012 stromal cocultures at concentrations of 0.1 ug/ml. Fusion protein
15 was replenished weekly in the stromal cocultures.

Competitive repopulating transplantation assay. Cultured cells were harvested, combined with fresh unfractionated BM obtained from congenic C57Bl/6 Ly5.1 mice (National Cancer Institute) and transplanted into lethally irradiated (10 Gy, split dose 3 h apart from a ¹³⁷Cs source, 1 Gy/min) Ly5.1 recipient mice. Each mouse received 2 X 10⁵
20 competitor BM cells and a fraction of the cocultured stem cells. Mice were bled by capillary puncture of the orbital venous plexus and 100 ul was collected; red blood cells were removed by NH₄Cl lysis. The nucleated cells were stained for the Ly5.2 (CD45.2) allelic marker using either fluorescein isothiocyanate-labeled directly conjugated Ly5.2 monoclonal antibody or a biotinylated form developed with streptavidin conjugated to Texas red. Cells
25 also were stained with directly conjugated antibodies to lineage markers. All antibodies and chromogens were obtained from Pharmingen. Flow cytometric analysis was done on a Becton Dickinson FACS Vantage using Cell Quest software.

Results

Genes expressed in AFT024 but not in 2018 were identified by a subtractive cloning
30 approach. Sequence analysis identified one of these AFT024-specific clones as dlk. Expression studies, (Figure 19A) show high levels of dlk in AFT024 and subclones isolated from this line, but undetectable levels in 2018 and BFC012. The latter two stromal cell lines

do not support repopulating stem cells. The PA6 stromal cell line and NIH Swiss 3T3 cells both show expression of dlk and were the cell sources for identification of SCP-1 and dlk, respectively. PA6 cells have been shown to support *in vitro* hematopoiesis and long-term, *in vivo* repopulating stem cells (Kodama, H., et al., 1992). Swiss 3T3 cells are also capable of supporting multipotent hematopoietic stem cells *in vitro*, promoting CSAs and maintaining *in vivo* spleen colony-forming units (CFU-S) (Roberts, R. A., et al., 1987). Interestingly, an additional fetal liver stromal cell line, 2012, which has some degree of stem cell supporting activity (Wineman, J., et al., 1996), and its subclones also express dlk. Furthermore, an RT-PCR analysis (40 cycles) of an additional 10 fetal liver-derived stromal cell lines and several other lines, shows detectable levels of dlk in only two additional lines (Figure 19B). These two cell lines (CFC032 and CFC008) can maintain some level of long-term transplantable stem cell activity present in whole BM (Wineman, J., et al., 1996). A correlation between a stromal cell line's ability to support stem cells and the expression of dlk was suggested. Therefore, functional studies were undertaken in order to delineate if dlk can act on or modulate hematopoietic stem cells.

Soluble dlk protein was added to progenitor cultures in semi-solid media. The soluble protein consisted of the dlk extracellular domain fused to the Fc portion of human IgG₁. The stem cell sources in these assays were highly enriched fetal liver cells (AA4.1⁺, lin^{lo/-}, Sca-1⁺, c-kit⁺). The influence of soluble dlk on hematopoietic progenitor colony-formation was assessed. As shown in Table 6, no differences were noted either in the number, sizes, or lineage compositions of colonies. Identical results were obtained at dlk concentrations ranging from 0.1 to 1.0 ug/ml. In addition, no differences were noted in similar studies using enriched BM cells (Sca-1⁺, c-kit⁺, lin^{-lo}).

Evidence for a positive effect of the dlk protein on stem/progenitor cells was observed when the soluble form was added to dexter-type cocultures. For these studies a stromal cell line (BFC012) was used that neither expresses endogenous dlk (see Figure 19) nor maintains significant *in vitro* hematopoiesis. In four experiments, two each using highly enriched adult BM and fetal liver stem cells, we monitored the appearance of CSAs over time. These colonies provide a convenient, quantitative estimate of hematopoietic activity initiated by primitive stromal dependent stem/progenitor cells. As shown in Figure 20, the addition of soluble dlk (0.1 ug/mL) results in an approximately 2-fold increase in the number of CSAs initiated by purified fetal liver or BM stem cells over a 2 week time period (P= 0.001 for dlk vs control and P= 0.01 for dlk vs no additive, Student's t-test). There was no

difference in the numbers of CSA observed in BFC012/stem cell cultures with or without control fusion protein (mean of no additive/control = 0.96 ± 0.11)

In order to assess the activity of the normal transmembrane form of dlk, a full-length dlk cDNA was transfected into BFC012 cells. Expression of the introduced dlk was demonstrated at the RNA (Northern blot) and protein levels using both Western blot and flow cytometric analyses with rabbit anti-dlk antibodies. The flow cytometry data are presented in Figure 21.

Initially, dlk-expressing transfected populations (BFC-dlk) were compared to a negative control "mock" transfected population of BFC012 cells. A 4- to 6-fold increase in the number of CSAs was observed in two separate experiments. The maintenance of CSAs was transient, lasting less than 2 weeks. No further hematopoietic activity was observed during an additional 2 weeks of culture. Once dlk-expressing clones had been identified from the transfected populations, they were studied for their ability to support CSAs in experiments designed to more precisely identify the time course of hematopoietic activity. Ficoll-separated, lineage depleted BM was used in these experiments. Five different negative control, non-dlk expressing BFC012 cell groups (parental BFC012 cells, two "mock" transfected populations, and two "mock" transfected clones) and 3 dlk-expressing BFC012 cell groups (one transfected population and two clones) were studied. The data are presented in Figure 22A. Neither the negative control BFC populations nor the "mock" transfected BFC clones supported high numbers of CSAs. In contrast, the BFC-dlk populations and the two individual dlk-expressing clones supported significantly greater numbers of CSAs at all time points studied ($P < 0.001$ days 3, 4, and 5; $P < 0.01$ days 6 and 7, Student's *t*-test). As observed previously, all the CSAs were transient. This experiment also indicated that the dlk-promoted hematopoietic activity peaks early, at 4 days, in this culture system. Three additional experiments using purified (AA4.1⁺, lin^{lo/-}, Sca-1⁺, c-kit⁺) fetal liver stem cells were performed using two individual clones, BFC-dlk-5 and a "mock" transfected negative control BFC-Zeo-1. The results are presented in Figure 22B. There was a dramatic and significant difference in the number of CSAs observed in the BFC-dlk-5 cultures compared to the control line ($P < 0.001$, days 4, 6, and 8, Student's *t*-test). As before, the effect was transient and the CSA declined in number over 2 weeks. AFT024 was included as a positive control and, in each of the three experiments, verified the quality of the input purified stem cells. In the first week of culture the numbers of CSAs observed on AFT024 were similar to the numbers in the BFC-dlk5 cultures.

In order to address the “primitiveness” of the cells that give rise to the CSA observed in the BFC-dlk cocultures, a series of *in vitro* replating experiments were performed. Individual wells were harvested at various time points of coculture and the cells were plated into semisolid cytokine-containing media. The numbers and lineage compositions of the colonies were scored after 8-12 days. As shown in Figure 23A, the CSAs obtained from day 4 BFC-dlk-5 cocultures contained numerous progenitors capable of extensive proliferation and multilineage differentiation. The total number of progenitors from the dlk-expressing cultures at day 4 was significantly expanded compared to the content in the freshly purified uncultured stem cell population ($P=0.01$, Student's *t*-test). The number and lineage composition of colonies derived from parallel day 4 AFT024 cultures was nearly identical to BFC-dlk-5 derived colonies. The content of CSAs replated at day 6 from the BFC-dlk-5 cocultures was devoid of multilineage colonies, although CFU-granulocyte-macrophages were maintained at high levels; the progenitor content in the BFC-dlk-5 cocultures continued to decrease when next sampled at 10 days. In contrast, few progenitors could be demonstrated in the BFC-Zeo-1 cultures ($P=0.001$, BFC-dlk-5 vs. BFC-Zeo-1, Student's *t*-test) (Figure 23A). Taken together, the data strongly suggest that dlk acts to promote stromal-dependent colony-formation by primitive cells capable of yielding large numbers of committed progenitors, including those endowed with a high proliferative capacity and multilineage differentiation potential. The lack of CSAs and significant progenitor maintenance in the BFC-Zeo-1 cultures argues that expression of dlk in the transfected BFC012 cells is responsible for both their ability to support CSAs and to generate/maintain primitive progenitors.

In order to determine if the CSA-containing cultures supported by BFC-dlk-5 contained stem cells capable of *in vivo* engraftment, portions of the same day 4 cocultures that were plated into progenitor assays also were used to transplant mice in competitive repopulation assays. Shown in Figure 23B are the results from two independent experiments analyzed at 10 weeks after transplant. The same BFC-dlk-5 cultures that contain CSAs and primitive CFU-high proliferative potential (HPP)-Mix progenitors also contain repopulating stem cells at levels equal to those maintained in parallel AFT024 cocultures. In addition, a significant difference exists in the levels of repopulating stem cells derived from dlk-expressing cocultures compared to non-dlk expressing BFC012 cells ($P=0.05$, Student's *t*-test). Multi-color flow cytometric analyses also demonstrated that both myeloid and lymphoid Ly5.2 cells are present in these animals. A subsequent analysis of these animals at

22 weeks demonstrated lower levels of repopulation with Ly5.2 cells derived from the AFT024 and BFC-dlk-5 supported cultures (data not shown). Most significantly, no repopulation was observed at any time point in mice that received cells cocultured on the non-dlk-expressing monolayers (Figure 23B).

5 Discussion

As part of an ongoing effort to understand the biology of the hematopoietic microenvironment, a panel of stromal cell lines from midgestation fetal liver was established and characterized. Among these cell lines, the AFT024 line has the ability to maintain nearly quantitative levels of transplantable stem cell activity for extended *in vitro* time periods (Moore, K. A., Ema, H. & Lemischka, I. R., 1997). Because these cultures are initiated with highly purified stem cell populations it is likely that the AFT024-derived molecular mechanisms responsible for this ability act directly on the stem cell population. Other stromal lines that fail to maintain stem cell activity were also identified. These observations facilitated a subtractive hybridization approach aimed at identifying potential candidate molecules whose collective actions may be responsible for the AFT024 stem cell maintenance activity. This effort has identified dlk, a transmembrane molecule containing six EGF-like repeat motifs. Although lacking the DSL motif indicative of the notch ligands delta and serrate (Tax, F. E., Yeagers, J. J. & Thomas, J. H., 1994), dlk is most closely homologous to delta/notch/serrate when compared to other EGF-like repeat containing molecules (Laborda, J., et al., 1993; Smas, C. M. & Sul, H. S., 1993). The predominant role of these types of molecules in cell growth and differentiation led us to investigate the potential role of dlk in hematopoiesis. Constitutive expression of translocated human notch (Tan-1) is found in a T-cell leukemia (Ellisen, L. W., 1991). Moreover the expression of Tan-1 in primitive human stem cells has been demonstrated (Milner, L. A., et al., 1994). Nevertheless, a functional role in hematopoiesis for the notch ligands Jagged (Lindsell, C. E., 1995) and Delta-like-1 (Bettenhausen, B., et al., 1995) has not been described. dlk expression is highly restricted in a panel of stromal cell lines. Two lines, AFT024 and 2012, which maintain repopulating stem cell activity *in vitro*, express dlk, whereas two non-supportive cell lines, 2018 and BFC012, do not. Interestingly, the S17 stromal cell line which is considered to be a potent stem cell supporter (Wineman, J. P., et al., 1993) does not express detectable levels of dlk. The S17 cell line was derived from adult BM (Collins, L. S. & Dorshkind, K., 1987)). The other lines described are all derived from fetal sources (AFT024, 2012, and NIH 3T3 cells) or from newborn calvaria (PA6 cells). It is therefore

possible that dlk acts in a developmentally regulated fashion. An extensive analysis of dlk expression in adult BM stroma is currently underway. Taken together, the data suggest that, at least in fetal stromal cell types, there exists a correlation between hematopoietic supporting ability and the expression of dlk.

5 The potential activity of both soluble and transmembrane dlk protein, on highly purified stem cell populations were directly measured using *in vitro* and *in vivo* assays. Initial experiments designed to ask if dlk can enhance colony formation in cytokine-rich semisolid assay systems were negative. These results may indicate that: (i) progenitor cells capable of colony formation in semisolid assays do not respond to dlk, (ii) the
10 collection of cytokines present in the semisolid cultures may "mask" an effect(s) of added dlk or (iii) that the soluble form of dlk requires a stromal monolayer to mediate its effects. The first possibility can be addressed more extensively in delta-type assays, where stem/progenitor cells are first cultured in suspension in serum-free media containing various cytokine combinations, with and without dlk, and then replated into colony assays (Muench,
15 J. O., Firpo, M. T. & Moore, M. A., 1993). The second possibility can be addressed by more extensive studies using subsets of the cytokines present in our initial studies. These experiments are underway. As a first step to address the third possibility, we added soluble dlk to preestablished BFC012 monolayers. Using both purified BM and fetal liver stem cell populations, a significant increase in CSA colony formation was observed in the dlk
20 supplemented cultures (Figure 20). This was a surprising result, given that dlk is a transmembrane protein; however, before its cDNA cloning, a soluble form of dlk was identified as FA1 or fetal antigen 1 (Jensen, C. H., et al., 1994). A role in hematopoiesis was not indicated in these studies, but expression was detected in stroma of placental villi, in yolk sac blood islands and in fetal liver (Jensen, C. H., et al., 1994). It is of interest to
25 determine if a soluble form is produced by the stromal cell lines that express dlk. An additional explanation, for the effects observed with the soluble form added to stromal/stem cell cocultures, is that they may be facilitated by the Fc portion of the fusion protein. It is possible that Fc receptors expressed by some of the hematopoietic cells in the cultures are able to sequester and present the dlk-Fc fusion protein more effectively. This possibility can
30 be addressed by using a different type of soluble dlk protein. These studies have been initiated. Alternatively, the soluble dlk may be sequestered and thus presented by the stromal cell extracellular matrix. In order to further address the third possibility, an intact transmembrane form of dlk was introduced into the BFC012 stromal cell line. Initially, dlk

transfected BFC012 cell populations were compared to BFC012 cells transfected with the selectable marker alone. In these studies, the dlk expressing monolayers were more effective at promoting CSA colonies. As with the soluble dlk experiments, the CSAs appeared early and were transient. When cells were replated from the BFC-dlk supported cocultures onto an AFT024 monolayer a reiteration of the burst of CSAs was seen that was maintained for 3 weeks. In addition, replating of these cocultures revealed a high progenitor content (~1 in 10) that included a high percentage (43%) of multipotential colonies including HPP-Mix. In the BFC-Zeo control populations neither replatable CSA nor CFU progenitors were maintained. Further experiments with individual clones from the transfected populations confirmed and extended the results obtained with the populations, demonstrating highly significant differences in the numbers of developing CSAs (Figure 22). However, in two experiments we observed that one dlk-transfected BFC012 clone, which expresses a very high level of dlk, supported fewer CSAs than non-dlk-expressing control cells. These cultures also suggested differentiation phenomena, as indicated by the number of rapidly accumulating nonadherent cells. Experiments utilizing this cell line were not included in our analyses. It is possible that there may be a threshold level of dlk expression necessary in these cultures and when it is surpassed the cells differentiate and die rapidly in the culture milieu provided by BFC012 cells. In addition, it is possible that an aberrant form of the dlk protein is made by this line. Further studies are necessary to clarify this issue. Nevertheless, in Figure 22A, dlk-transfected BFC012 cells (one population and two clones) show a significant enhancement of CSA formation compared to controls.

The observed low level maintenance of competitive repopulating stem cells in short-term dlk-expressing cocultures is of interest even though the activity diminished over time. These studies show that the ectopic expression of a single molecule (dlk) in a previously nonsupporting stromal cell line restores or enables hematopoietic support. This is demonstrated by maintenance of three different stem/progenitor cell compartments: (i) CFU-HPP-Mix, (ii) CSA, and (iii) short-term *in vivo* repopulating stem cells. It is also of interest that both qualitatively and quantitatively similar stem/progenitor cell compartments are maintained in short-term AFT024 supported cocultures (Moore, K. A., Ema, H. & Lemischka, I. R., 1997).

Two mechanisms underlying the effects of dlk are considered. First, it may be that some level of dlk expression is sufficient to retard potent differentiation signals provided by the BFC012 cell line. Second, dlk may provide a proliferative stimulus not normally

produced by BFC012. A more direct and perhaps relevant assay will be to eliminate the expression of dlk in AFT024 cocultures, thus maintaining other components that make up the culture "milieu." Studies to evaluate potential neutralizing antibodies and various dlk antisense strategies are underway.

5 The failure of BFC-dlk-5 supported CSAs to persist for periods longer than 1 to 2 weeks also may suggest the existence of other molecules in AFT024 that facilitate hematopoiesis. In this regard, it is interesting that our subtraction screen has yielded several other clones with expression patterns very similar to dlk. Eventually, with the addition of dlk and other AFT024 specific molecules it may be possible to reconstruct a
10 supportive phenotype. This should lead towards a further understanding of the *in vivo* hematopoietic microenvironment. In summary, it is proposed that dlk represents one molecular component responsible for the hematopoietic supportive activities of the AFT024 cell line. As such, dlk may define a novel molecular pathway of stem cell regulation by the hematopoietic microenvironment.

15

REFERENCES

The following is a list of documents related to the above disclosure and particularly to
20 the experimental procedures and discussions. The documents should be considered as incorporated by reference in their entirety.

1. Andrews, R., J. Singer, and I. Bernstein. 1989. J Exp Med 169:172 1.
- 25 2. Ardman, B., et al. (1992). Proc. Natl. Acad. Sci. U S A. 89, 5001-5005.
3. Artavanis-Tsakonas, D., K. Matsuno, and M.E. Fortini. 1995. Science 268:225-232.
4. Ball TC, Hirayama F, Ogawa M. Blood 85:3086, 1995
- 30 5. Barbosa, M.D., Q.A. Nguyen, V.T. Tchernev, J.A. Ashley, J.C. Detter, S.M. Blalysdes, S.J. et al., 1996. Nature 382:262-265.
6. Baud, W., et al., 1995. Genomics 26:334-344.
7. Bauer, D., et al., Nucl Acids Res 21:4272-4280.
8. Baum, C.M., et al., Proc Natl Acad Sci USA 89:2804-2808.

9. Bauman, J.G., et al., *Acta Histochem* 36:241-253.
10. Bazil, V., Brandt, et al. (1997). *Stem Cells* 15, Suppl. 1, 13-18.
- 5 11. Beavis, A.J., and K.J. Pennline. 1994. *J Immunol Methods* 170:57-65.
12. Bettenhausen, B., et al., 1995 *Development* 121, 2407-2418.
13. Bhatia, M., et al., 1997. *Proc Natl Acad Sci USA* 94,5320-5325.
- 10 14. Bonnet, D., and J.E. Dick. 1997. *Nature Med* 3:730-737.
15. Bork P, et al. *Q Rev Biophys* 1996 May;29:119-167
16. Bornemann, D., E. Miller, and J. Simon. 1996. *Development* 122:162 1 1630.
17. Borrow, J., et al., 1996. *Nat Genet* 12:159-167.
- 15 18. Bradford, G.B., et al., 1997. *Exp Hematol* 25:445-453.
19. Brady, G., and N.N. Iscove. 1993. In *Methods in Enzymology*, 225: 611-623.
20. Brady, G., et al. 1990. *Methods in Molecular and Cellular Biology* 2:17-25.
21. Brandt J, et al., *J Clin Invest* 86:932, 1990
- 20 22. Braun, B.S., et al., 1995. *Mol Cell Biol*, no. Aug., 1995:4623-4630.
23. Britten, R.J., et al., 1985. *Hybridisation Strategy*. In *Nucleic Acid Hybridisation*.
B.D. Homes and S.J. Higgins, editors. Oxford, Washington, DC. 3-15.
24. Broxmeyer HE, et al. *Proc Natl Acad Sci USA* 89:4109, 1992
- 25 25. Brugger W, et al. *Blood* 81:2579, 1993
26. Carpino, N., et al., 1997. *Cell* 88:197-204.
27. Cashman, J., et al., 1997. *Blood* 89:4307-4316.
- 30 28. Cheng, J., et al. (1996). *Blood* 87, 479-490.
29. Cheng, L., et al., 1996. *Nature Biotechnology* 14:606.
30. Chiang, M.-Y., et al. *J Biol Chem* 266:18162-18171.

31. Chomczynski, P., and Sacchi, N. (1987). *Anal. Biochem.* 162, 156-159.
32. Chubet, R.G., and B.L. Brizzard. 1996. *BioTechniques* 20:136-141.
33. Collins LS and Dorshkind K: *J Immunol* 138:1082, 1987
34. Collins, L. S. & Dorshkind, K. (1987) *J. Immunol.* 138, 1082-1087.
- 5 35. Craig, W., et al., 1993 *J Exp Med* 177:1331-1342.
36. Cumano, A., et al., (1993). *Proc. Natl. Acad. Sci. USA.* 90, 6429-6433.

37. Deiss, L., et al., 1995. *Genes Dev* 9:15-30.

- 10 38. Deryugina, E. I., et al., 1994 *Exp. Hematol.* 22, 910-918.

39. Dexter TM, in Wright DG, Greenberger JS (eds): *Long term bone marrow culture.*
New York, Alan R. Liss, Inc., 1984, p 57
40. Dexter, T. M., et al., (1977) *J. Cell. Physiol.* 9, 335-344.
- 15 41. Diatchenko, L., et al. 1996. *Proc Natl Acad Sci USA* 93:6025-6030.
42. DiGiusto, D.L., et al., 1996 *Blood* 87:1261-1271.
43. Dorshkind K *Annu Rev Immunol* 8:11, 1990
44. Dosil, M., et al., 1996. *Blood* 88, no. 12:4510-4525.
45. Elledge, S., and J. Harper. 1994. *Curr Opin Cell Biol* 6:847-852.
- 20 46. Ellisen, L. W., et al., (1991) *Cell* 66, 649-661.

47. Fairbairn, L.J., et al., 1993. *Cell* 74:823-832.

48. Fleming, W.H., et al., 1993 *J Cell Biol* 122:897-902.
- 25 49. Frampton, J., et al., 1995 *EMBO J* 14:2866-2875.
50. Fraser CC, et al., *Blood* 76:1071, 1990
51. Frederiksen K, et al., *Neuron* 1:439, 1988
52. Friedrich, C., et al., (1996) *Blood* 87, 4596-4606.
53. Fukai, K., et al., 1996 *Am J Hum Genet* 59:620-624.
- 30 54. Galy, A., et al., *Immunity* 3:459-473.
55. Goode, S., et al., 1996 *Development* 122:3863-3879.
56. Goode, S., et al., 1992 *Development* 116:177-192.
57. Goodwin, R. G., (1990) *Cell* 60, 941-951.

58. Graf, L., and B. Torok-Storb. 1995. Blood 86:548-556.
59. Guan, E., et al. J. Biol. Chem. 266, 20345-20355.
60. Guan, E., et al. (1994). J. Immunol. 152, 4005-4012.
61. Gubler, U., and Hoffman, B.J. (1983).
- 5 62. Hamann, J., et al., 1995. J Immunol 155:1942-1950.
63. Hao Q-L, et al., Blood 88:3306, 1996
64. Hao, Q.L., et al., 1995. Blood 86:3745-3753.
65. Hao, Q.L., et al., 1996. Blood 88:3306-3313.
- 10 66. Harper, J., and S. Elledge. 1996. Curr Opin Genet Dev 6:56-64.
67. Harrison DE, et al., Exp Hematol 21:206, 1993
- 15 68. Harrison DE, et al., Blood 69:1021, 1987
69. Harrison DE, Stone M, Astle CL. J Exp Med 172:431, 1990
70. Harrison, D.E. 1980. Blood 55:78.
71. Harrison, S.M., et al., 1995 Development 121:2479-2489.
- 20 72. Hawley, R., et al., 1997. Cancer Res 57:337-345.
73. Hawley, R., et al., 1994. Gene Ther 1: 136-138.
74. Heximer, S., et al., 1996 DNA Cell Biol 15: 1025-1038.
75. Hofmann, A., et al., 1996 Proc Natl Acad Sci USA 93, no. May:5185-5190.
- 25 76. Holyoake TL, et al., Blood 87:4589, 1996
77. Hopp, T. P., et al., (1988) Biotechnology 6, 1204-1210.
78. Horan, PK., et al., 1990 Methods Cell Biol 33:469-490.
- 30 79. Hu, M., 1997 Genes Dev 11:774-785.
80. Huang S, et al., Blood 83:1515, 1994

81. Hubank, M., and D.G. Schatz. 1994. Nucl Acids Res 22, no. 25:5640-5648.
82. Issad C, et al., Blood 81:2916, 1993
- 5 83. Jensen, C. H., et al., (1994) Eur. J. Biochem. 225, 83-92.
84. Jones, R., et al., 1996. Blood 88:487-491.
85. Jones, R., et al., 1990. Nature 347:188-189.
- 10 86. Jongens, T.A., et al., 1992 Cell 70:569-584.
87. Jordan, C. T., et al., 1995. Exp. Hematol. 23, 1011-1015.
88. Jordan, C.T., and Lemischka, I.R. (1990). Genes Dev. 4, 220-232.
- 15 89. Jordan, C.T., et al. Cell 61, 953-963.
90. Jordan, C.T., et al. 1996. Exp Hematol 24:1347-1355.
91. Kabrun, N., et al. (1997). Development 124, 2039-2048.
92. Keller, G.M. (1995). Cell. Biol. 7, 862-869.
- 20 93. Khochbin, S., D. et al., 1990 Cytometry II: 869-874.
94. Kinsella, T.M., and Nolan, G.P. (1996). Hum Gene Ther. 12, 1405-1413.
95. Kitamura, T., et al. Proc. Natl. Acad. Sci. U S A. 92, 9146-9150.
96. Knobel KM, et al., Exp Hematol 22:1227, 1994
97. Knobel, K. M., et al., (1994) Exp. Hematol. 22, 1227-1235.
- 25 98. Ko, M.S.H. 1990. Nucl.Acids Res 18:5705-5711.
99. Kodama H, et al., J Exp Med 176:351, 1992
100. Kodama H, et al., J Cell Physiol 118:233, 1984
101. Koller MR, et al., Blood 82:378, 1993
102. Kulesa, H., et al., 1995 Genes Dev 9:1250-1262.
- 30 103. Laborda, J., et al., 1993 J. Biol. Chem. 268, 3817-3820.
104. Lansdorp PM, et al., J Exp Med 175:1501, 1992
105. Lapidot-Lifson, Y, et al., 1992 Proc Natl Acad Sci USA 89:579-583.

106. Larochelle, A., et al., 1997 *Nature Med* 2:1329-1337.
107. Lawrence, H.J., et al., 1996 *Stem Cells* 14:281-291.
- 5 108. Lemischka IR: *Dev Biol* 4:379, 1993
109. Lemischka IR, *Cancer Surv* 15:3, 1992
110. Lemischka, I.R. (1997). *Blood* 89, 4337-4347.
111. Lemischka, I.R. 1992. *Curr Top Micro Immunol* 177:59-71.
112. Li SH, et al. *J Biol Chem* 1994 269:4431-4437
- 10 113. Li, C.L., and G.R. Johnson. 1992. *J Exp Med* 175:1443.
114. Li, W.-B., et al., 1994. *BioTechniques* 16:722-729.
115. Li, X., et al., 1995. *Cytometry* 20:172-180.
116. Liang, P., et al., 1994 In *Methods in Molecular Genetics*, 5: 3-16
117. Lida, S., et al., 1993 *Oncogene* 8:3085-3092.
- 15 118. Lindsell, C. E., (1995) *Cell* 80, 909-917.
119. Ling, V., and Neben, S. (1997). *J Cell Phys.* 171, 104-115.
120. Lord BI, et al., *Blood* 45:65, 1975
121. Lowry, P.A., et al., 1995 *Exp Hematol* 23:1135-1140.
- 20 122. Lyons, A.B., et al., 1992 *Cytometry* 13:809-821.
123. Mackarechtschian, K., 1995 *Immun* 3:147-161.
124. Marcos, M.A.R. et al. (1997). *J. Immunol.* 158, 2627-2637.
125. *Methods in Mol Biol* 5:303-322.
126. Matthews, W., et al., 1991 *Cell* 65:1143-1152.
- 25 127. Matthews, W., 1991 *Proc Natl Acad Sci USA* 88:9026-9030.
128. Mayani, H., 1993 *Blood* 82:2664-2672.
129. McKearn, J.P., et al. (1985). *Proc. Natl. Acad. Sci. U S A.* 82, 7414-7418.
130. Mehta P, et al. 1997 90:2381-2389
- 30 131. Metcalf D: *The molecular control of blood cells*. Cambridge, MA, Harvard University Press, 1988
132. Milner, L. A., et al., (1994) *Blood* 83, 2057-2062.
133. Milner, L.A., 1996 *Proc Natl Acad Sci USA* 93:13014-13019.

134. Mishell BB, et al.,: Selected Methods in Cellular Immunology. San Francisco, CA, W. H. Freeman and Company, 1980, p 23
135. Moore et al., Br J Haematol 18:279, 1970
- 5 136. Moore, K.A., et al., 1997 Proc Natl Acad Sci USA 94:4011-4015.
137. Moore, K.A., H. Ema, and I.R. Lemischka. 1997. Blood 89:4337-4347.
138. Morrison, L.E., et al. (1991). Oncogene 6, 1657-1665.
139. Morrison, S.J., et al., 1997 Development 124:1929-1939.
- 10 140. Morrison, S.J., and I.L. Weissman. 1994. Immunity 1:661-673.
141. Morrison, S.J., et al., 1995 Ann Rev of Cell Dev Biol 11:35-71.
142. Morse HC, et al., Immunogenetics 25:71, 1987
143. Muench MO, et al., Leuk Lymphoma 16:1, 1994
144. Muench, J. O., et al., (1993) Blood 81, 3463-3473.
- 15 145. Mulder AH, et al., Exp Hematol 15:99, 1987
146. Muller, C., et al., EMBO J 14:61276135.
147. Murray, L., et al., 1994. Blood Cells 20:364-369.
148. Nakamura, T., et al., 1996 Nat Genet 12:154-158.
- 20 149. Neben S, et al., Exp Hematol 21:438, 1993
150. Neering, S., et al., 1996 Blood 88:1147-1155.
151. Nepomuceno, R.R., et al. (1997). Immunity 6, 119-129.
- 25 152. Neuman-Silberberg, F.S., and T. Schupback. 1996 Mech Dev 59:105-113.
153. Nolte, J.A., et al., 1996. Proc Natl Acad Sci USA 93:2414-2419.
154. Ogawa, M. (1993) Blood 81, 2844-2853.
- 30 155. Okabe, T., et al., 1995.. J Immunol 154:38713879.
156. Okada, S., et al., (1992) Blood 80, 3044-3050.
157. Okuda, T., et al., 1996 Cell 84:321-330.

158. Orkin, S.H. 1996. *Curr Opin Genet Dev* 6:597-602.
159. Osawa, M., et al., *Science* 273:242-245.
160. Patanjali, S.R., et al., 1991 *Proc Natl Acad Sci USA* 88:1943 -1947.
161. Pawliuk, R., et al., 1996 *Blood* 88:2852-2858.
- 5 162. Perou, C.M., et al., 1996 *Nat Genet* 13:303-308.
163. Persons, D., et al., 1997 *Blood* 90:1777-1786.
164. Peters SO, et al., *Exp Hematol* 23:461, 1995
165. Peters, S. O., et al., (1995) *Exp. Hematol.* 23, 461-469.
166. Petrenko, O., (1997). *Oncogene* 14, 1671-1680.
- 10 167. Petzer, A.L., et al., 1996. *Proc Natl Acad Sci USA* 93:1470-1474.
168. Ploemacher RE, et al., *Blood* 74:2755, 1989
169. Ploemacher, R. E., et al., (1991) *Blood* 78, 2527-2533.
170. Ploemacher, R.E., et al., 1989 *Blood* 74:2755.
171. Porcher, C., et al., 1996 *Cell* 86:47-57.
- 15 172. Randall, T.D., et al., 1996 *Blood* 87:4057-4067.
173. Rebay, I., et al., (1991) *Cell* 67, 687-699.
174. Rebel VI, et al., *Blood* 83:128, 1994
175. Robb, L., et al., 1996 *EMBO J* 15:4123-4129.
176. Roberts, R. A., et al., (1987) *J. Cell Physiol.* 132, 203-214.
- 20 177. Ropp, J., et al., 1995 *Cytometry* 21:309-317.
178. Roth, S., et al., 1995. *Cell* 81:967-978.
179. Sambrook, J., et al., (1989) *Molecular Cloning: A Laboratory Manual* (Lab. Press, Plainview, New York).
- 25 180. Sauvageau, G., U. et al., 1995 *Genes Dev* 9:1753-1765.
181. Schmid, I., et al., 1994 *Cytometry* 15:12-20.
182. Shalaby, F., et al., 1995. *Nature* 376:62-66.
183. Shaw, G., and R. Kamen. 1986. *Cell* 46:659-667.
- 30 184. Sherr, C. 1993. *Cell* 73:1059.
185. Sherr, C. 1996. *Science* 274:1672-1677.
186. Shivdasani, R.A., et al., *Blood* 87:4025-4039.
187. Shockett, P.E., et al., 1996. *Proc Natl Acad Sci USA* 93:5173-5176.

188. Sieweke, M., et al., 1996 Cell 84:49-60.
189. Simon, J. 1995. (1993) Cell 73, 725-734.
190. Soares, M.B., et al., 1994 Proc Natl Acad Sci USA 91:9228-9232.
- 5 191. Soma T, et al., Blood 87:4561, 1996
192. Spangrude GJ, Science 241:58, 1988
193. Spangrude, G.J., et al., 1990, Proc Natl Acad Sci USA 87:7433.
194. Spangrude, G.J., et al., 1995 Blood 85:1006-1016.
- 10 195. Spangrude, G.J., 1988 Science 241:58.
196. Stockton, B.M., et al. (1998). Immunity 8, 373-381.
- 15 197. Sutherland, H.S., et al., 1989. Blood 74:1563.
198. Suzuki J, et al., Leukemia 6:452, 1992
199. Suzuki, A., et al. (1996). Blood 87, 3550-3562.
200. Szilvassy SJ, et al. Blood 87:4618, 1996
201. Taswell C, J Immunol 126:1614, 1981
- 20 202. Tax, F. E., Yeagers, J. J. & Thomas, J. H. (1994) Nature 368, 150-154.
203. Telford, W.G., L.E. King, and P.J. Fraker. 1992. Cytometry 13:137-143.
204. Telford, W.G., L.E. King, and P.J. Fraker. 1994. J Imm. Meth 172 1-16.
205. Terstappen LWMM, et al. Blood 77, 1991
- 25 206. Testa NG, Molineux G: Haemopoiesis: A practical approach. New York, New York, Oxford University Press, Inc., 1993
207. Traycoff CM, et al., . Exp Hematol 24:299, 1996
208. Traycoff, C., et al., 1994. Exp Hematol 22, no. 215.
- 30 209. Traycoff, C., et al., 1994 Prog Clin Biol Res 389:281-29 1.
210. Trentin JJ: Regulation of Hematopoiesis, vol. 1, 1970
211. Trevisan, M., and N.N. Iscove. 1995. J Exp Med 181:93-103.
212. Trevisan, M., X.-Q. Yan, and N.N. Iscove. 1996. Blood 88:4149-4158.

213. Tsai, S., et al. (1994). *Genes Dev.* 23, 2831-2841.
214. Tsai, S., et al., 1994. *Genes Dev* 8:2831-2841.
215. Tsai, S., et al., 1992 *Genes Dev* 6:2258-2269.
216. Tu L, et al. 1996 *J Immunol* Nov 1;157:3995-4004
- 5 217. Turner, C., et al., 1996. *Blood* 87:3237-3244.
218. Uchida, N., et al., 1995 In *Hematopoietic Stem Cells: Biology and Therapeutic Applications*. Merteismann, editor.
219. Uchida, N., et al., 1993 *Curr Opin Immunol* 5:177-184.
220. Uckun, F. 1990. *Blood* 76:1908-1923.
- 10 221. Van der Sluijs JP, et al., *Leukemia* 7:725, 1993
222. Verfaillie CM, and Miller JS: *Blood* 86:2137, 1995
223. Verfaillie, C., et al., 1990 *J Exp Med* 172:509-520.
224. Visvader, J., 1992 *EMBO J* 11:4557-4564.
225. Wang, Q., 1996. *Proc Natl Acad Sci USA* 93:3444-3449.
- 15 226. Wang, Z., and D.D. Brown. 1991. *Proc Natl Acad Sci USA* 88:11505-11509.
227. Watson SR, et al. *J Cell Biol* 1991 115:235-243
228. Weilbaecher, K., et al., 1991. *Blood* 78:945-952.
229. Weiss L, et al., *Blood* 78:975, 1991
230. Whitlock CA, and Witte ON *Proc Natl Acad Sci USA* 79:3608, 1982
- 20 231. Whitlock, C.A., and C.E. Müller-Sieburg. 1990.
232. Wineman, J. P., et al., (1993) *Blood* 81, 365-372.
233. Wineman, J., et al., (1996) *Blood* 87, 4082-4090.
234. Wolf, N. S. (1979) *Clin. Hematol.* 8, 469-500.
235. Wood, H.B., et al. *Blood* 90, 2300-2311.
- 25 236. Yamanashi, Y., and D. Baltimore. 1997. *Cell* 88:205-211.
237. Yang, E., and S. Korsmeyer. 1996. *Blood* 88:386-401.
238. Yang, T., L. Cheng, and S. Kain. 1996. *Nucl Acids Res* 24:4592.
239. Yang, Y., et al., 1996. *Exp Hematol* 24:605-612.
- 30 240. Yoneda-Kato, N., et al., 1996 *Oncogene* 18, no. 12(2):265-275.
241. Yonemura, Y., 1996 *Proceedings Natl Acad Sci USA* 93:4040-4044.

242. Yonemura, Y., et al., 1997 Blood in press.

243. Young, B.D., and M.L.M. Anderson. 1985. In Nucleic Acid Hybridisation. B.D. Homes and S.J. Higgins, editors. Oxford, Washington, DC. 47-71.

5 244. Young, J.C., et al., 1996 Blood 87, no. 2:545-556.

245. Zettlmeissl, G., et al., (1990) DNA Cell Biol. 9, 347-353.

246. Zhang, L., et al., 1997. Science 276:1268-1272.

10 This invention may be embodied in other forms or carried out in other ways without departing from the spirit or essential characteristics thereof. The present disclosure is therefore to be considered as in all respects illustrative and not restrictive, the scope of the invention being indicated by the appended Claims, and all changes which come within the meaning and range of equivalency are intended to be embraced therein.

15

WHAT IS CLAIMED IS:

1. An isolated nucleic acid derived from an isolated hematopoietic stem cell, the isolated nucleic acid comprising the following characteristics:
 - 5 - specifically expressed in the hematopoietic stem cell; and
 - encoding a hematopoietic stem cell – specific protein.
2. An isolated nucleic acid of claim 1, the isolated nucleic acid further comprising the following characteristic:
 - 10 - capable of hybridizing under standard conditions with a sequence selected from the group consisting of SEQ.ID.No.: 1, SEQ.ID.No.: 2, SEQ.ID.No.: 3, SEQ.ID.No.: 4, SEQ.ID.No.: 5, SEQ.ID.No.: 6, SEQ.ID.No.: 7, SEQ.ID.No.: 8, SEQ.ID.No.: 9, SEQ.ID.No.: 10, SEQ.ID.No.: 11, SEQ.ID.No.: 12, SEQ.ID.No.: 13, SEQ.ID.No.: 14, SEQ.ID.No.: 15, SEQ.ID.No.: 16, SEQ.ID.No.: 17, SEQ.ID.No.: 18, SEQ.ID.No.: 19, SEQ.ID.No.: 20, SEQ.ID.No.: 21, SEQ.ID.No.: 22, SEQ.ID.No.: 23, SEQ.ID.No.: 24, SEQ.ID.No.: 25, SEQ.ID.No.: 26, SEQ.ID.No.: 27, SEQ.ID.No.: 28, SEQ.ID.No.: 29, SEQ.ID.No.: 30, SEQ.ID.No.: 31, SEQ.ID.No.: 32, SEQ.ID.No.: 33, SEQ.ID.No.: 34, SEQ.ID.No.: 35, SEQ.ID.No.: 36, SEQ.ID.No.: 37, SEQ.ID.No.: 38, SEQ.ID.No.: 39, SEQ.ID.No.: 40, SEQ.ID.No.: 41, SEQ.ID.No.: 42, and SEQ.ID.No.: 43, SEQ.ID.No.: 45, SEQ.ID.No.: 47, SEQ.ID.No.: 49, SEQ.ID.No.: 51, SEQ.ID.No.: 53, SEQ.ID.No.: 55, SEQ.ID.No.: 57, SEQ.ID.No.: 59, SEQ.ID.No.: 61, SEQ.ID.No.: 63, SEQ.ID.No.: 65, SEQ.ID.No.: 67, SEQ.ID.No.: 72 or a portion thereof.
- 25 3. The isolated nucleic acid of claim 1, the isolated nucleic acid further comprising the characteristic of encoding a protein capable of modulating hematopoietic stem cell activity.
- 30 4. The isolated nucleic acid of claim 3, wherein the activity is selected from the group consisting of hematopoietic stem cell differentiation and hematopoietic stem cell replication.

5. The isolated nucleic acid of claim 3, wherein the protein is selected from the group consisting of a growth factor, a transcription factor, a splicing factor, a capping factor, a transport protein, a translation factor, and a replication factor
- 5 6. The isolated nucleic acid of claim 1, wherein the nucleic acid comprises the nucleotide sequence of SEQ.ID.No.: 72, an analog thereof, or a portion thereof.
7. The isolated nucleic acid of claim 1, wherein the hematopoietic stem cell is a primitive hematopoietic stem cell.
- 10 8. The isolated nucleic acid of claim 7, wherein the primitive hematopoietic stem cell is selected from the group consisting of an umbilical cord cell, a bone marrow cell and a fetal liver cell.
- 15 9. The isolated nucleic acid of claim 7, wherein the primitive hematopoietic stem cell is selected from the group consisting of a AFT024 cell, a 2012 cell and a 2018 cell.
10. A composition comprising the nucleic acid of claim 1, wherein the nucleic acid comprises one selected from the group consisting of SEQ.ID.No.: 1, SEQ.ID.No.: 2,
20 SEQ.ID.No.: 3, SEQ.ID.No.: 4, SEQ.ID.No.: 5, SEQ.ID.No.: 6, SEQ.ID.No.: 7, SEQ.ID.No.: 8, SEQ.ID.No.: 9, SEQ.ID.No.: 10, SEQ.ID.No.: 11, SEQ.ID.No.: 12, SEQ.ID.No.: 13, SEQ.ID.No.: 14, SEQ.ID.No.: 15, SEQ.ID.No.: 16, SEQ.ID.No.: 17, SEQ.ID.No.: 18, SEQ.ID.No.: 19, SEQ.ID.No.: 20, SEQ.ID.No.: 21, SEQ.ID.No.: 22, SEQ.ID.No.: 23, SEQ.ID.No.: 24, SEQ.ID.No.: 25, SEQ.ID.No.: 26, SEQ.ID.No.: 27,
25 SEQ.ID.No.: 28, SEQ.ID.No.: 29, SEQ.ID.No.: 30, SEQ.ID.No.: 31, SEQ.ID.No.: 32, SEQ.ID.No.: 33, SEQ.ID.No.: 34, SEQ.ID.No.: 35, SEQ.ID.No.: 36, SEQ.ID.No.: 37, SEQ.ID.No.: 38, SEQ.ID.No.: 39, SEQ.ID.No.: 40, SEQ.ID.No.: 41, SEQ.ID.No.: 42, SEQ.ID.No.: 43, SEQ.ID.No.: 45, SEQ.ID.No.: 47, SEQ.ID.No.: 49, SEQ.ID.No.: 51, SEQ.ID.No.: 53, SEQ.ID.No.: 55, SEQ.ID.No.: 57, SEQ.ID.No.: 59, SEQ.ID.No.: 61, SEQ.ID.No.: 63, SEQ.ID.No.: 65, SEQ.ID.No.: 67, SEQ.ID.No.: 72 or a portion thereof.
30

11. The nucleic acid of claim 1, wherein the nucleic acid is selected from the group consisting of DNA, RNA and cDNA.
12. A vector comprising the nucleic acid of claim 1.
- 5 13. The vector of claim 12, wherein the vector comprises viral or plasmid DNA.
14. An expression vector comprising the nucleic acid of claim 1 and a regulatory element.
- 10 15. A host vector system which comprises the expression vector of claim 12 in a suitable host.
16. The host vector system of claim 15, wherein the suitable host is selected from the group consisting of a bacterial cell, a eukaryotic cell, a mammalian cell and an insect cell.
- 15 17. An isolated hematopoietic stem cell specific protein or a portion thereof encoded by the nucleic acid of claim 1.
- 20 18. The isolated hematopoietic stem cell of claim 16, further comprising the following characteristic:
- capable of modulating hematopoietic stem cell activity.
19. The isolated hematopoietic stem cell of claim 18, wherein the activity is selected from
- 25 the group consisting of hematopoietic stem cell differentiation and hematopoietic stem cell replication.
20. The protein of claim 17, wherein the protein is selected from the group consisting of a growth factor, a transcription factor, a splicing factor, a capping factor, a transport protein, a translation factor, and a replication factor.
- 30 21. The protein of claim 17, the protein having substantially the same amino acid sequence as one selected from the group consisting of SEQ.ID.No.: 42, SEQ.ID.No.:

44, SEQ.ID.No.: 46, SEQ.ID.No.: 48, SEQ.ID.No.: 50, SEQ.ID.No.: 52, SEQ.ID.No.: 54, SEQ.ID.No.: 56, SEQ.ID.No.: 58, SEQ.ID.No.: 60, SEQ.ID.No.: 62, SEQ.ID.No.: 64, SEQ.ID.No.: 66, SEQ.ID.No.: 68, SEQ.ID.No.: 70, SEQ.ID.No.: 71, and SEQ.ID.No.: 73.

5

22. A nucleic acid probe capable of specifically hybridizing with the nucleic acid of claim 1 under standard conditions.

23. An antibody capable of specifically binding to the protein of claim 17 without substantially cross-reacting with a non-stem cell specific protein or homologs thereof under conditions permissive to antibody binding.

10

24. A cell capable of producing the antibody of claim 23.

25. A method for identifying the presence of a primitive hemopoietic stem cell in a sample comprising:

15

(a) contacting the sample with the antibody of claim 23 under conditions permissive to the formation of an antibody complex; and

20

(b) detecting the presence of the complex formed in step (a), the presence of a complex formed indicating the presence of a primitive hemopoietic stem cell in the sample.

26. The method of claim 25, wherein the antibody is labeled with a detectable marker.

27. The method of claim 26, wherein the detectable marker is selected from the group consisting of a radioactive isotope, enzyme, magnetic bead, dye, fluorescent marker and biotin.

25

28. A method for generating a stem cell/progenitor cell from a primitive hematopoietic cell in a sample comprising contacting the sample with the protein of claim 17.

30

29. A method for identifying the presence in a sample of a compound that modulates hematopoietic stem cell activity comprising:
- (a) contacting the hematopoietic stem cell with the sample;
 - (b) determining the hematopoietic stem cell activity;
 - 5 (c) comparing the hematopoietic stem cell activity determined in step (b) with the activity determined in the absence of the compound an increase or decrease in hematopoietic stem cell activity indicating the presence in the sample of a compound that modulates hematopoietic stem cell activity.
- 10 30. The method of claim 29, wherein the activity is selected from the group consisting of gene expression, replication, differentiation, transplantation, and self regeneration.
31. A compound identified by the method of claim 29, previously unknown.
- 15 32. A method for identifying primitive hematopoietic stem cell-specific nucleic acids , comprising:
- (a) creating a primitive hematopoietic stem cell cDNA library and a non-primitive stem cell immune cell cDNA library; and
 - (b) subtracting the two libraries, thereby identifying primitive stem cell specific
20 nucleic acids.
33. The method of claim 32, step (b) comprising:
- (i) contacting the nucleic acids of the stem cell and non-stem cell libraries with each other under conditions permissive to hybridization, thereby forming
25 hybrid complexes;
 - (ii) separating the hybrid complexes formed in step (b) from the nucleic acids which did not form complexes;
 - (iii) isolating the nucleic acids which did not form complexes, thereby identifying hematopoietic stem cell specific nucleic acids.
30
34. The method of claim 33, step (ii) further comprising amplification of the nucleic acids.

35. The method of claim 33, step (iii) further comprising amplification of the nucleic acids which did not form complexes.
36. The method of claim 33, further comprising displaying the amplified DNA on a chromatography gel.
37. The method of claim 32, step (b) comprising differential display of the two libraries, thereby identifying primitive stem cell specific nucleic acids.
38. The method of claim 32, step (b) comprising representation difference analysis of the two libraries, thereby identifying primitive stem cell specific nucleic acids.
39. The method of claim 32 further comprising cloning the stem cell specific nucleic acid.
40. The method of claim 32, wherein the stem cell is selected from the group consisting of AF024, 2012, and 2018.
41. A nucleic acid identified by the method of claim 32
42. A composition comprising the compound of claim 31 and a carrier.
43. A pharmaceutical composition comprising the compound of claim 31 and a pharmaceutically acceptable carrier.
44. A primitive hematopoietic stem cell specifically expressing one selected from the group consisting of: SEQ.ID.No.: 1, SEQ.ID.No.: 2, SEQ.ID.No.: 3, SEQ.ID.No.: 4, SEQ.ID.No.: 5, SEQ.ID.No.: 6, SEQ.ID.No.: 7, SEQ.ID.No.: 8, SEQ.ID.No.: 9, SEQ.ID.No.: 10, SEQ.ID.No.: 11, SEQ.ID.No.: 12, SEQ.ID.No.: 13, SEQ.ID.No.: 14, SEQ.ID.No.: 15, SEQ.ID.No.: 16, SEQ.ID.No.: 17, SEQ.ID.No.: 18, SEQ.ID.No.: 19, SEQ.ID.No.: 20, SEQ.ID.No.: 21, SEQ.ID.No.: 22, SEQ.ID.No.: 23, SEQ.ID.No.: 24, SEQ.ID.No.: 25, SEQ.ID.No.: 26, SEQ.ID.No.: 27, SEQ.ID.No.: 28, SEQ.ID.No.: 29, SEQ.ID.No.: 30, SEQ.ID.No.: 31, SEQ.ID.No.: 32, SEQ.ID.No.: 33, SEQ.ID.No.: 34, SEQ.ID.No.: 35, SEQ.ID.No.: 36, SEQ.ID.No.: 37, SEQ.ID.No.: 38, SEQ.ID.No.: 39,

SEQ.ID.No.: 40, SEQ.ID.No.: 41, SEQ.ID.No.: 42, and SEQ.ID.No.: 43, SEQ.ID.No.: 45, SEQ.ID.No.: 47, SEQ.ID.No.: 49, SEQ.ID.No.: 51, SEQ.ID.No.: 53, SEQ.ID.No.: 55, SEQ.ID.No.: 57, SEQ.ID.No.: 59, SEQ.ID.No.: 61, SEQ.ID.No.: 63, SEQ.ID.No.: 65, SEQ.ID.No.: 67, SEQ.ID.No.: 72 or a functional portion thereof.

5

45. A primitive hematopoietic stem cell specifically expressing a nucleic acid identified by the method of claim 32.

46. A method for treating a condition in a subject comprising administering to the subject a therapeutically effective amount of the composition of claim 43.

10

47. The method of claim 46 wherein the condition is an immune system condition.

48. The method of claim 46 wherein the condition is leukemia.

15

49. A method of introducing an exogenous nucleic acid into a hematopoietic stem cell comprising contacting the stem cell with the composition of claim 42.

50. A method of ex vivo expansion of hematopoietic stem cells comprising contacting the cell with the composition of claim 42.

20

FIGURE 1

Figure 1. Subtracted Libraries

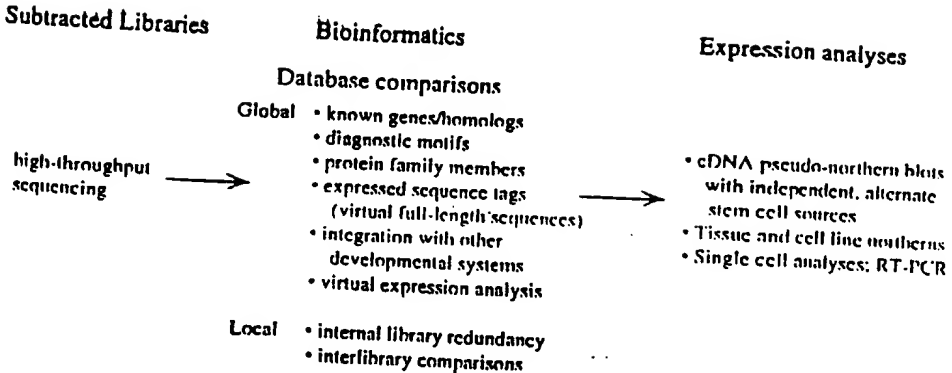


FIGURE 2

Figure 2.

Cyt-19

SA61

Smc-34

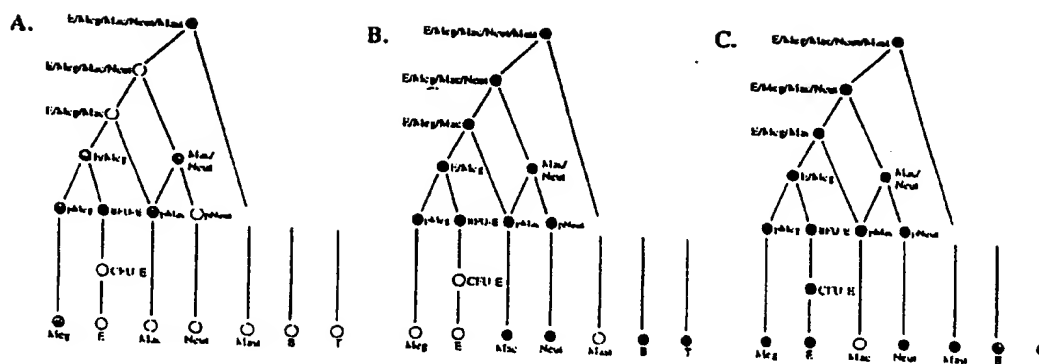


FIGURE 3

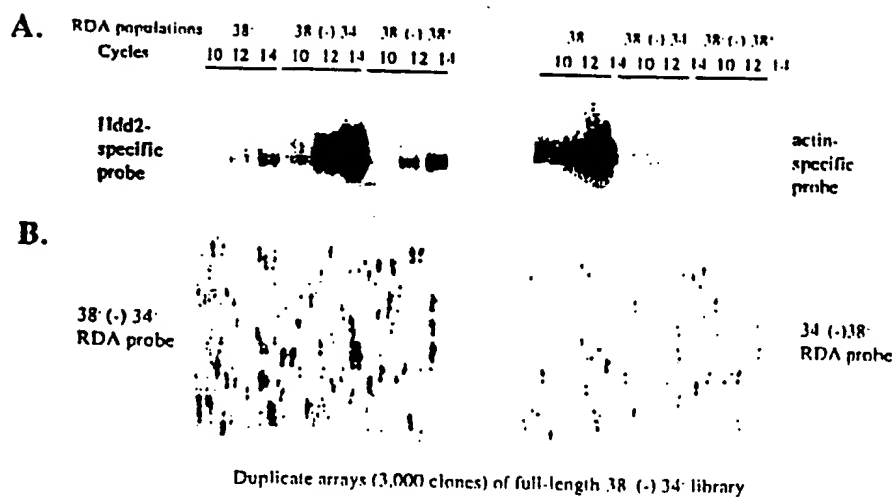


FIGURE 4

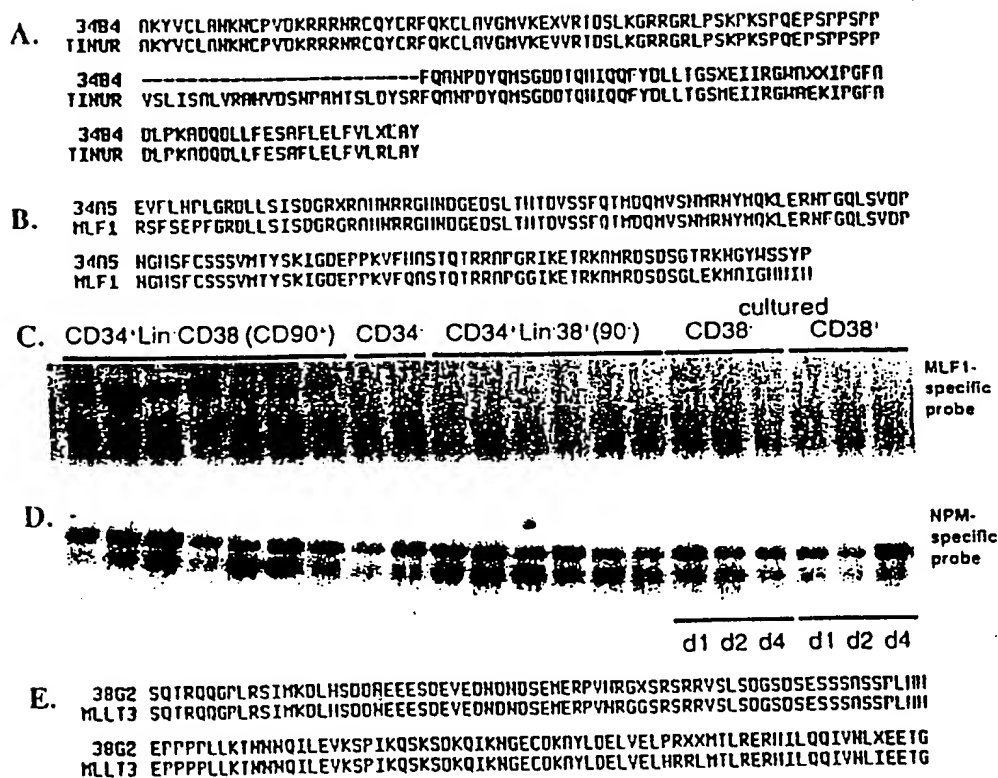
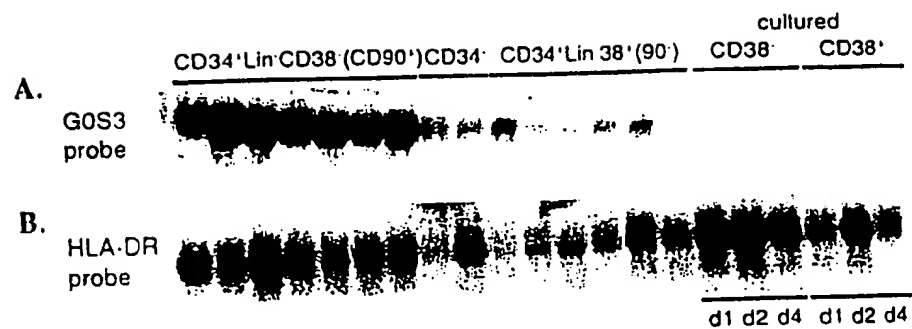


FIGURE 5

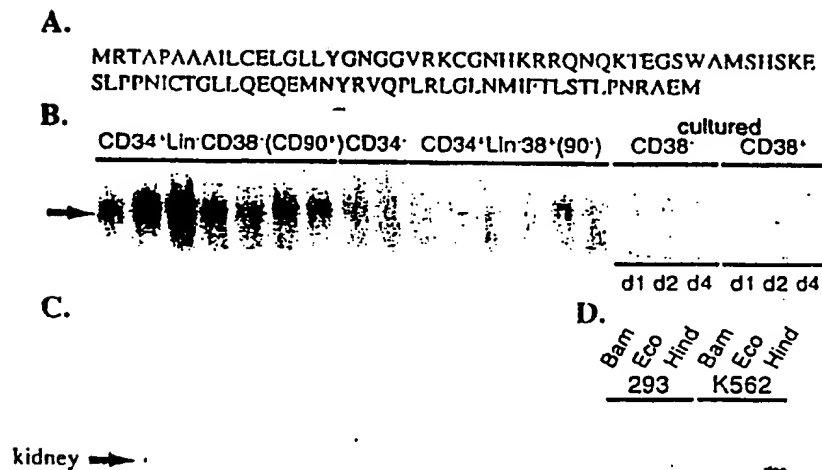
Figure 5.



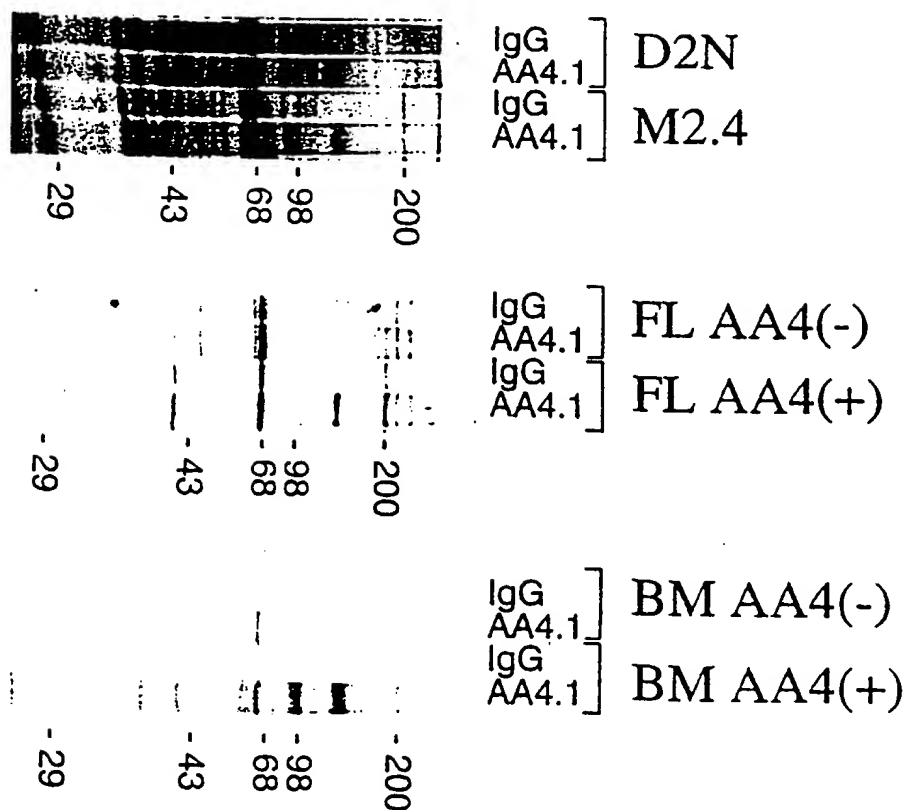
6/99

FIGURE 6

Figure 6.



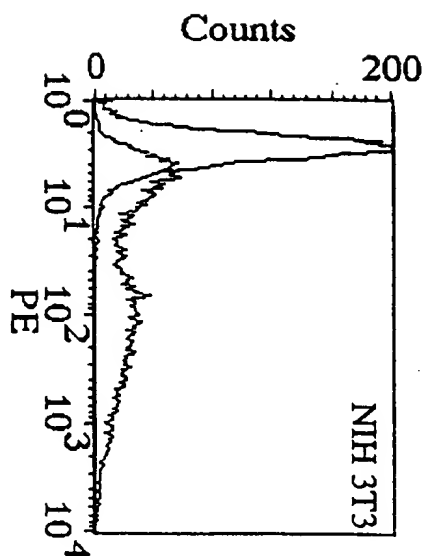
7/99

FIGURE 7

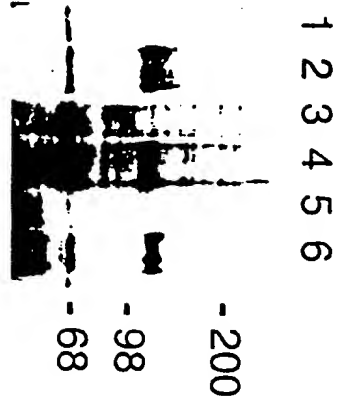
8/99

FIGURE 8

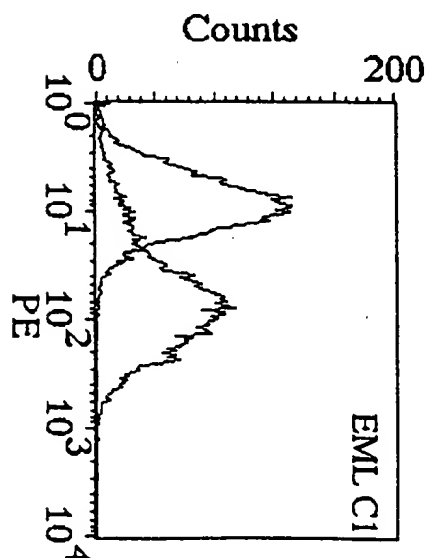
A



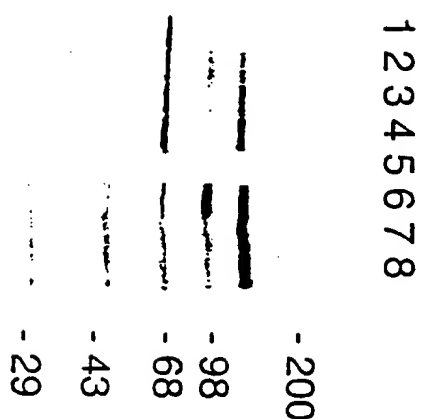
B



A



C



[illegible]

10/99

FIGURE 10

A

AA4 HAISTGLFLLGLL-GQWAGAAASQAVCEGTACTYTAHWGKLSAAEAQHONENQGNL 59
 MA S GL LLL LL QP AG AD--AWVC GTACTYTAH GKLSAAEAQ+ CN+NGGNL
 C1qR MATSMGLLLLLLLLLLTQPGAGTGADTEAVVCVGTACTYTAHSGKLSAAEAQHONENQGNL 60

AA4 ATVKSEEEARHVQOALTQLLTKAPLEAVNGKFWIGLQREKGNCTYHDLPHRGFSWVGQ 119
 ATVKS+EEA+HVQ+ L QLL+ +A L A+M KFWIGLQREK C LP++GFSWVGQ
 C1qR ATVKSKEEAQHVQRLAQLLRRAALTARMSKFWIGLQREKGLDPSLPLKGFWSVGQ 120

AA4 EDTAYSNNWYASKSSCIFKRCVSLILDLSLTPHPSHLPKWHESPCGTPEARQNSIEGFLC 179
 EDT YSNW+K ++SCI KRCVSL+LDLS P+ LPMW E PCG+P +PG++IEGF+C
 C1qR EDTPYSNWYKELRNSCISKRCVSLILDLSQPLLPNRLPKWSEGPCSGPGSGNSIEGFVC 180

AA4 KFNFKGHCRLALGGPGRVITYTTPFQATSSLEAVPFASVANVACGDEAKSET--HYFLC 237
 KF+FKGHCRLALGGPG+VTYTPFP T+SSLEAVPFAS ANVAC+ K ET HYFLC
 C1qR KFSFKGHCRLALGGPGQVITYTTPFQTSSLEAVPFASVANVACGEGDKDETQSHYFLC 240

AA4 NEKTPGIFHWSSGPLCVSPKFGCSFNNGCQDCFEQDGSFRCCRPGRLLDDLVT 297
 EX P +F WSSGPLCVSPK+GC+FNNGC QDCFEQDGSF CCRPGFRLLDDLVT
 C1qR KEKAPDVFHWSSGPLCVSPKYGCNFNNGCQDCFEQDGSFLCGCRPGFRLLDDLVT 300

AA4 ASRNPCSSNPCTGGGHCHSVPLSENITCRCPGQYQLDSSQVHCVDIDECQDSPCAQCQVN 357
 ASRNPCSS+PC GG C P +NYTCRCP GYQLDSSQ+ CVD+DECQDSPCAQ+CVN
 C1qR ASRNPCSSSPCRGGATCVLGPNGKNTCRCPGQYQLDSSQVHCVDIDECQDSPCAQCQVN 360

AA4 TLGSFHCCHWGYQPSGPKEEACEDVDCAAAANSPCAQCCINTDGSFYCSCKEGYIVSGE 417
 T G F CECWGY+P GP E AC+DVDECA SPCAQCC NTDGSF+CSC+EGY+++GE
 C1qR TPGGRFCEC+NGYEPGGPGEACQDVEDCALGRSPCAQCCNTDGSFYCSCKEGYIVSGE 420

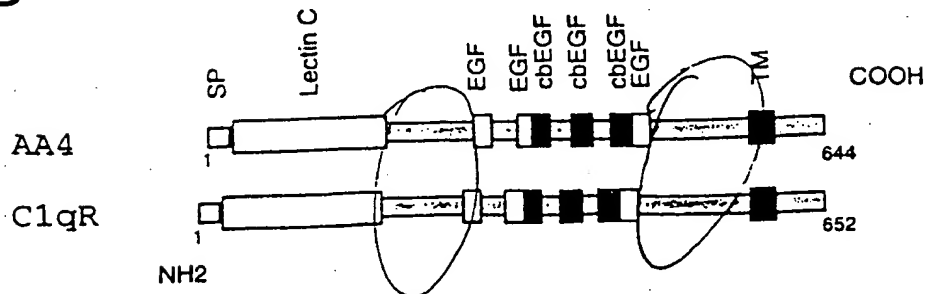
AA4 DSTQCEDIDECSDARGNPCDSLFCFNTDGSFRCCPQWELAPNGVFCRGTVFSPLPAP 477
 D TQC+D+DEC G CDSLFCFNT GSF CCC PGW LAPNGV C+ G V P+ P
 C1qR DSTQCQDVEDCVGPGPLCDSLFCFNTDGSFHCGLPGWELAPNGVSCITNGPVSLGPPSGP 480

AA4 PQKEDNDRKESTHPTTEMPSSPSGSKDVSNRAQTGL-FVQSDIPTASVPLEIIPSEV 536
 P +ED +++ ST+P S G+ TT +SD P S PL++ PS
 C1qR PDEEDKGEKSTVTPRAATASPTRGPEGTPKATPITSRPLSSDAPITSAPLKHLAPSGS 540

AA4 SDWVFELOTYLP-TSGHSPKTHEDSVSA---HSDTDQNNLLLFYILGTVAISLLLVLA 592
 S VW E + T SG +P DS A + TDQ LLLFYILGTVAI LLL LA
 C1qR SGWREPSIHATAASGPQEPAGGSSVATQNDGTGQKLLLFYILGTVAIILLALLA 600

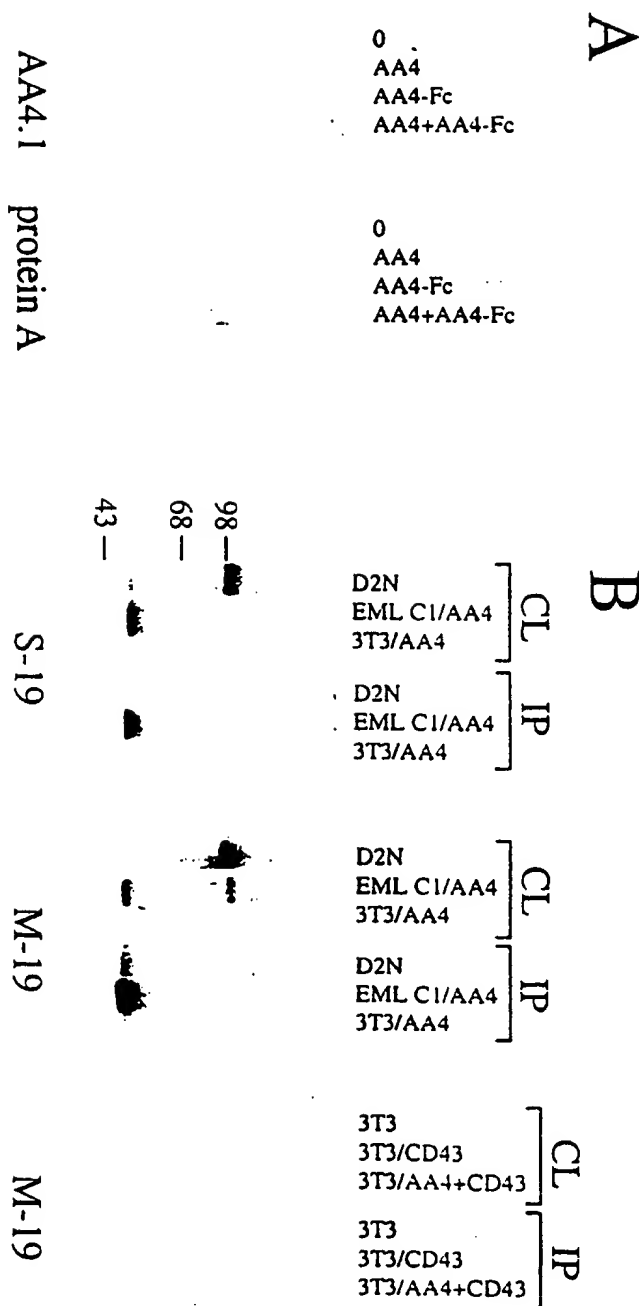
AA4 LGILYHKRRAKKEEIKKKPQNAADSYSWPERAESQAPENQYSPTPGTDC 644
 LG+L+Y KRRAK+EE KEKKPQNAADSYSWPERAES+A ENQYSPTPGTDC
 C1qR LGLLVYRKRRAKKEEKKKKPQNAADSYSWPERAESRAMENQYSPTPGTDC 652

B



11/99

FIGURE 11



12/99

FIGURE 12



13/99

FIGURE 13

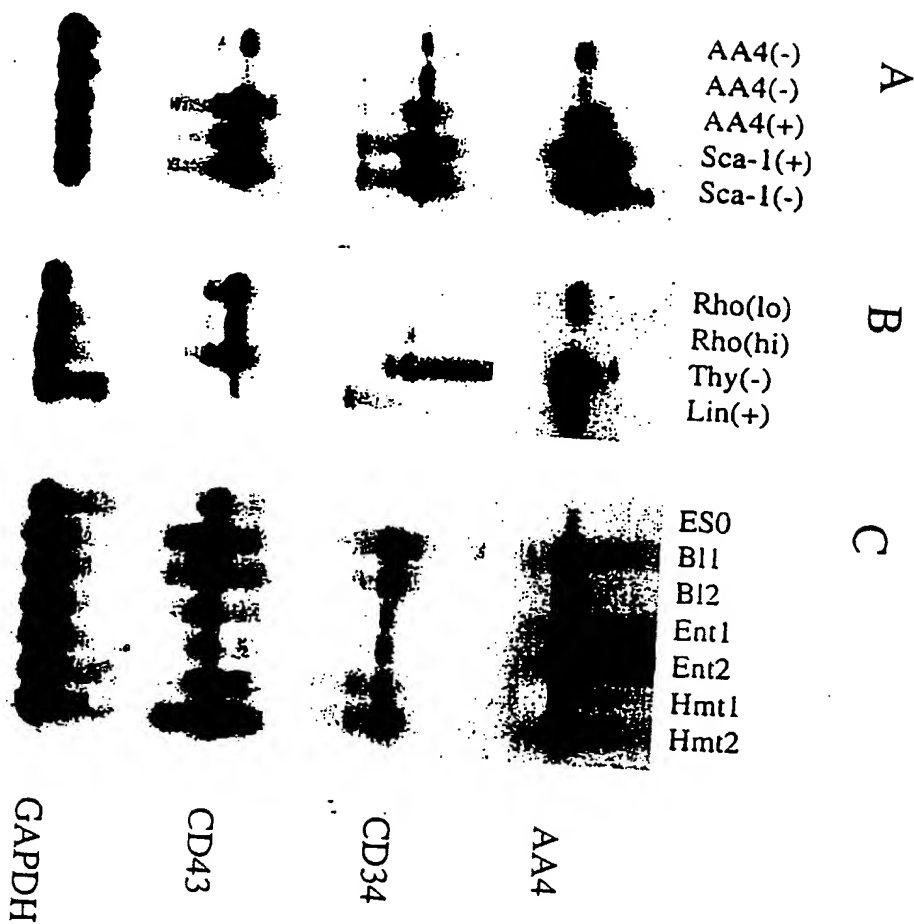
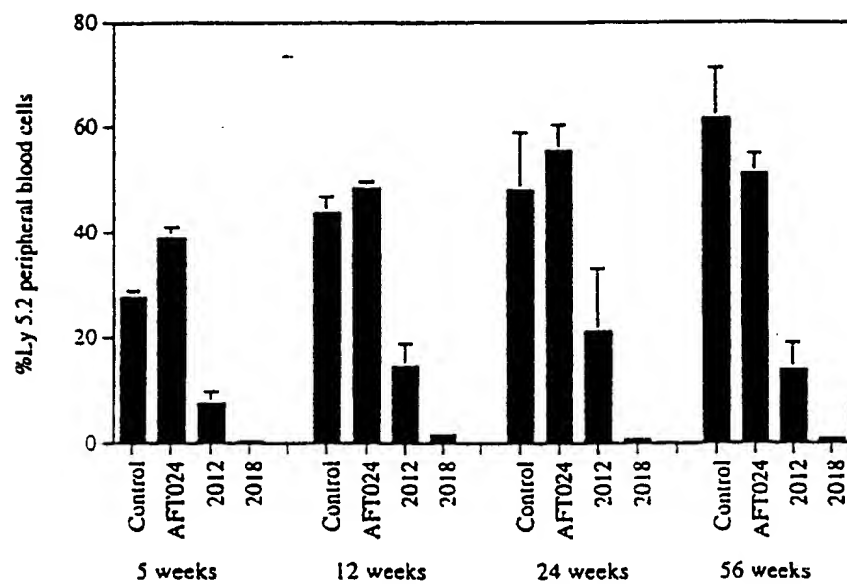
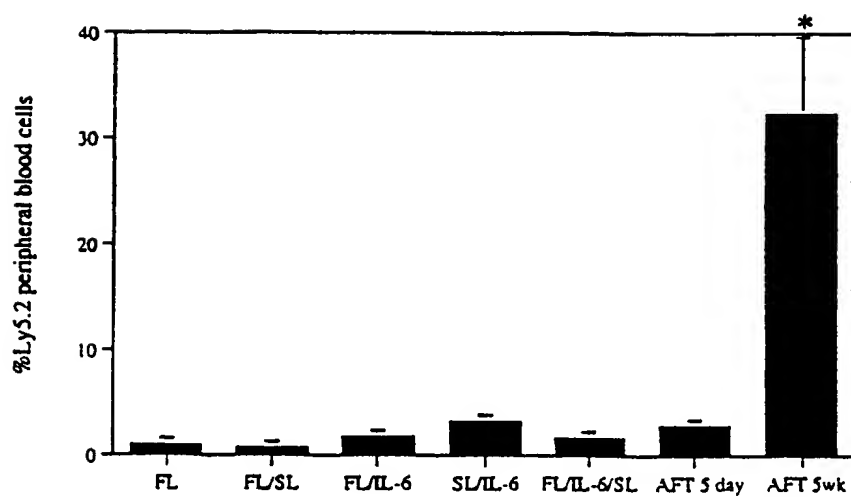


FIGURE 14

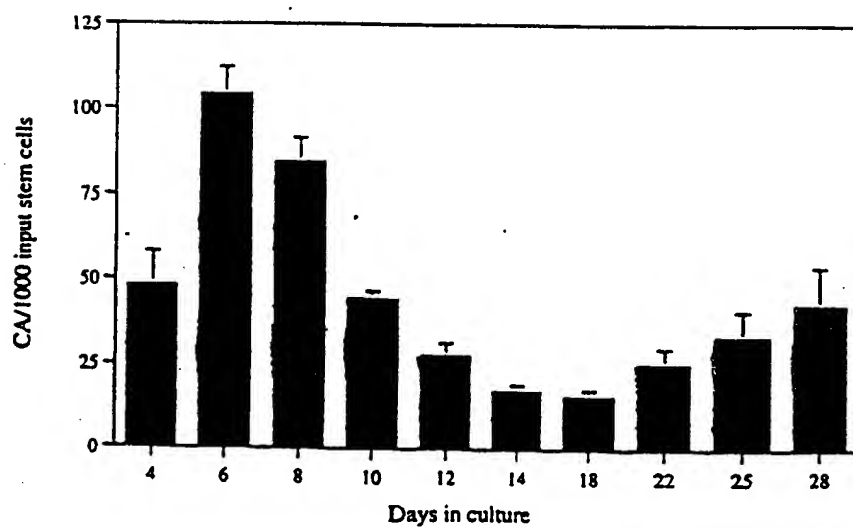
15/99

FIGURE 15



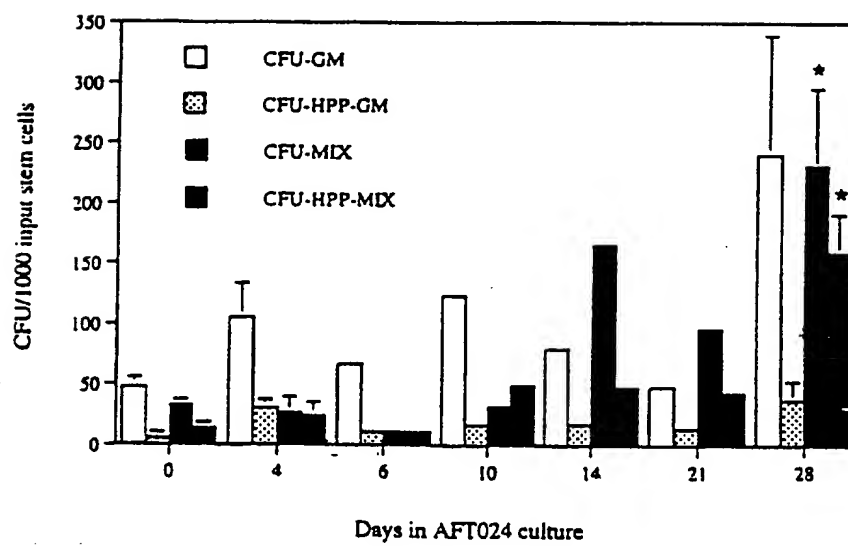
16/99

FIGURE 16

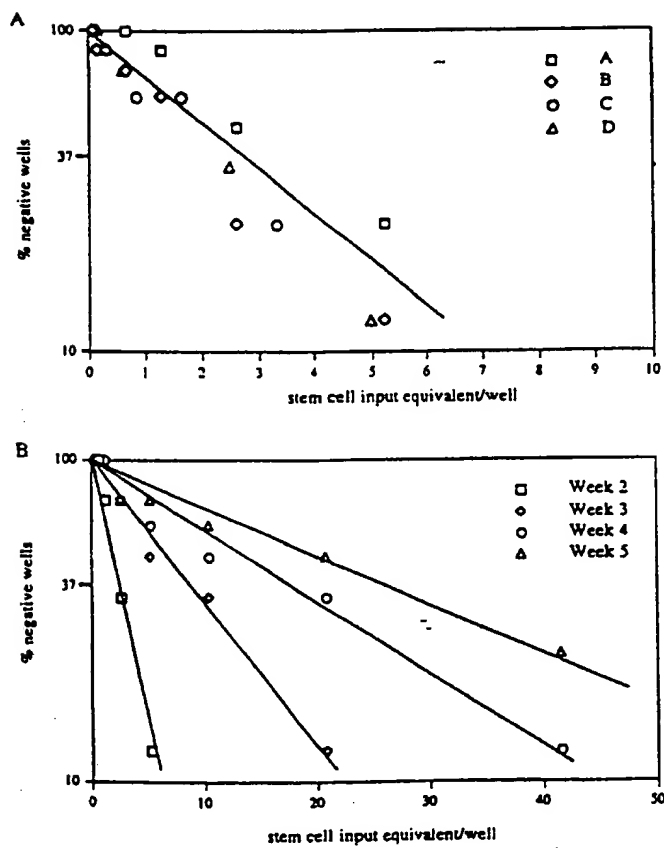


17/99

FIGURE 17

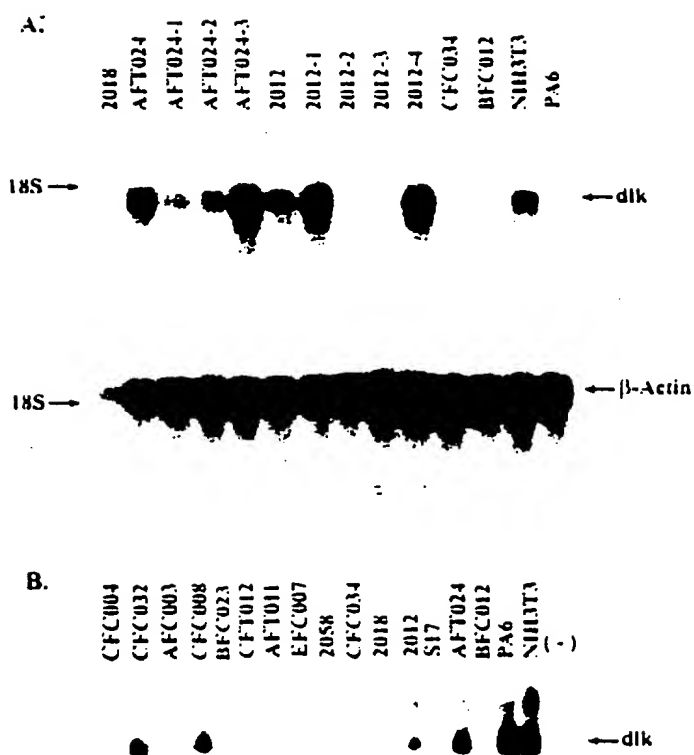


18/99

FIGURE 18

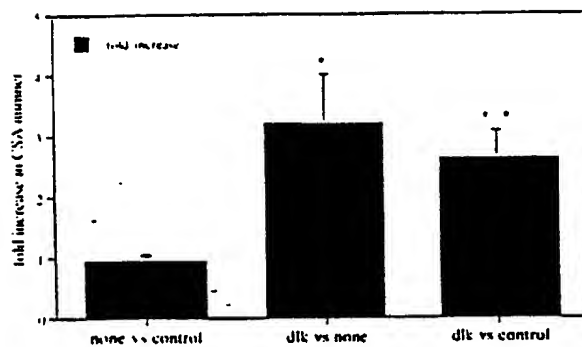
19/99

FIGURE 19



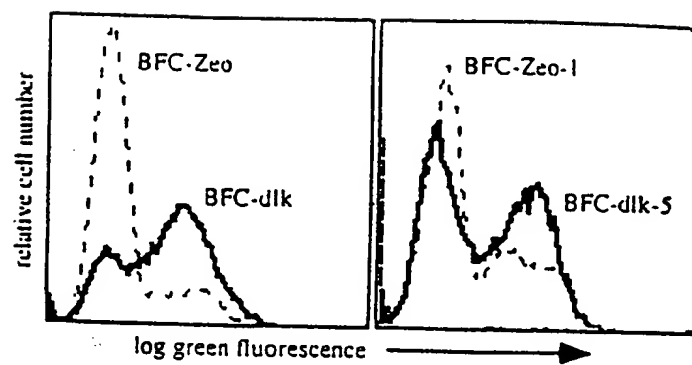
20/99

FIGURE 20



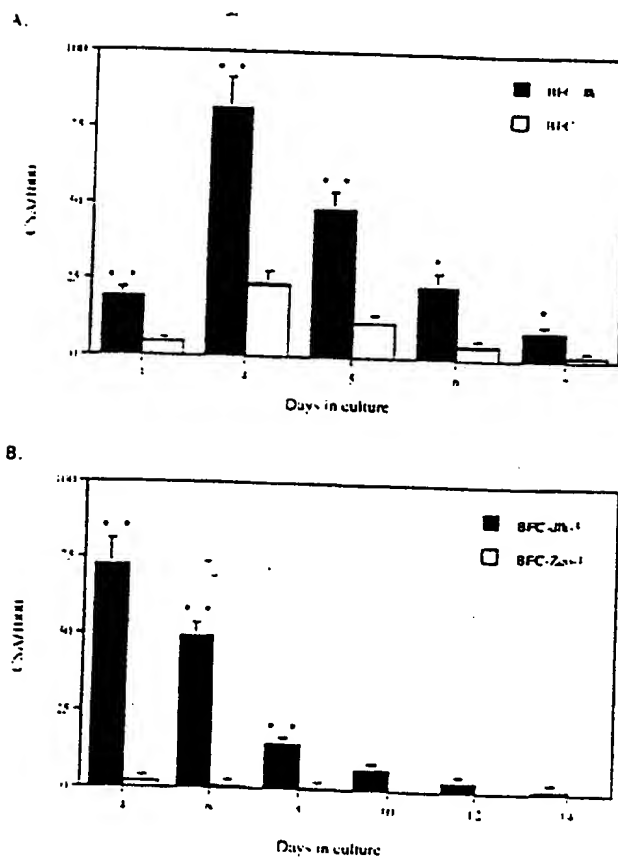
21/99

FIGURE 21



22/99

FIGURE 22



24/99

LL4-39 (GL34)

Session Name: crick.princeton.edu 4

Page 1

1 CCGAGGCGT: CCGAGAGCGG CAGCCATCTC CATGCCAAGG TTGAACMAAA
51 TCCAGCAAA: GTTCTTTTCT GCTAGCCCA AGGATTGGA AGAACACAGC
101 CATTCTTT: TAGCGACAA AAACAGAAA AAAGGCCAA AGACTTTTGG
151 AATGACGTT: AAACATACC TGAGATGAT GATCCACAT CTGGAATCTG
201 GATGAAAT: GCGCAATCC AAACACATC TTCTCTCTGA AGAAGTATG
251 CAGTGGTCT: AGTCTCTGA AAATCTCTG CCAACAGAC AGGTGAAAT
301 GTCTTGAA: GATTCMAAG TGTATTGAG TGAGGAAAT ATTGAATCT
351 GGTGGCTT: TGAGACTAT KAKAAACAN AAATGATCT TTTCATANC

SEQ ID. No.: 1

FIGURE 24

25/99

23.41 (0.54)

Session Name: crick.princeton.edu 4

Page 1

```
1. CCGCTGCGAG AAGAGGACAG AAGGGGACTC ACTGTTTTT AGAAGACCAT
5. GAGAGCGGCA GCGTCTCCA TCCAGGTT GACAAATG CCAGGAATGT
101 TCTTTCTGC TATCCAAAG GATTCGAAN AACACAGCA TTCTCTTC
```

SEQ ID. No.: 2

FIGURE 24

26/99

Bl-66 (related to G05B)

Session Name: crick.princeton.edu 4

Page 1

```
1  TACGGCTGCC AGAAGACGAC AGAAGGGGAA NCAGCTGCGG TGGCGCGCGG
91  AGTCTGACAA TCGAAGTGC CATGTTCTTG GCTGTCGAC ACGACTGCGT
101 ACCCATGGAC AAGAGTGCAN GCAACGGCCC CAGGTCHEN GAGAAGCGG
151 AGAAATGAJ. TCGACACTC TTTGANGATT GGAAGACCGG TTTCAGCTAC
201 TTCTGCGAGJ. GAGCTCTGTC TCGCGGAGG CCGTGAAGTG GCAATAAAC
251 NAGAGCATA GTTTEATCA NCCTTCTGCT NANGAGGCG ANCTCTGCG
301 ATAACTTT: NANGAGTGC TGGCGGTAT ATATGGGCTG GCTGATTCA
351 NCGCTTTT: TAAGTCCA GTTCTGAA GAAACATG AATCTGCT
```

SEQ ID. No.: 3

FIGURE 24

27/99

LL2-12 (b2, b6, b7C)

Session Name: crick.princeton.edu 1

Page 1

```
1  CCTGTGCTG OCTTCGACC CCTTCCTTC CTGCTCATCC TGTGTGTAC
51  GGGGCGAG CACCTGACC AGAGAACGC CATTCGGCA TCTTGGGTG
101 CCATCGCGA AGCGGGGT 'TTCAGATCC AGACGCTCTT CCTCTGGA
151 AACTTGA A CACACACT TCTGACTG TCTCAGAT CACACACA
201 CAGGAGATC TTGAGGCT CCTTCAGGA TTCTACGC AACCTACCC
251 TCAAGACT CATGACTG AACTGGTGA ACAAATCTG TCTATGCC
301 CCTACATC TCAAGCGA TATGACTG TATGTAAG TCCAGACT
351 GTTTCAGG CTGATACAA GAGGGGGCC TTGAGACT GGCAGAGG
```

SEQ ID. No.: 4

FIGURE 24

uLS-96 (112) obl

Session Name: crick.princeton.edu 4

Page 1

```
1  CCGAGGCGTC CGATCTCTTC CAGGCGCCAC AAGCAGCTCT GAAGAGCCAT
51  GTTCGAGCT: GCGGAGCCG CCGAGGCCAC CCGCTCTCAT GAAGCCAAAG
101 GCGGCG/TC: CAGCAGCAGT GTACAGCGGT CTAAGTCTTT TACCTTCCG
151 CCTCAGCTGA AAGAGAGCTG TCGAGCTTC CAGAGAGCTG TGTACCCGAT
201 GGAGCGGCT: GTGCGAGACA AGCTCATTTT CCACAACTCT TTTTCTGTT
251 GCAAGAGCT: CCACGCCAAA CTCAGGCTGG GCACTTATTC TGCATGCGC
301 GGTGATTTT: ACTGCGAGCC TCACTTTTCA GAGCTTTTA AGATTAAAG
351 CAACTACCA: GAGGTTTGG TGTAAACAG CACAAGGACT CTGCGCCAC
```

SEQ ID. No.: 5

FIGURE 24

29/99

U.S.-96 (line) 35

Session Name: crick.princeton.edu 1

Page 1

```
1  ATGTGTTACG: CTGACGGGAC CGGTCCGGAA TTCCCGGTC GACCAAGCGT
51  CCGCCGACG: GTCCGATCTC CTCCAGGGG CACCAAGGAC CTCTGAAGAA
101 CATCTTCA  GCTGCCGAG (CTCCGAGC CACCCCTCT CNTGAAGGCC
151 AAGGCMKAC TGGCAGCAG ACCGTNCAOC GGTCTAAGTC CTTTAACTTG
201 CCGGCTCAG TGAAGGAGAC CTGTGAGCC TGGCGAAGA CTGTGTNCC
251 NATGGAGCG CTGTGAGCAG ACAAGTCTT TTTCACAGC TCTGTCTCT
301 GTTGCAGAC CTGCCACAC AAGTCAACC TGGCAGTTA TGTGTCTAT
351 CACGTTAAT TTTACTGAG AGCTACTTT CAGCAGCTGT TTAGAATNA
401 AGGCGATAC NATGAGAGT CTGTCTCTA ACAGCAGAG GAGCTCTGG
451 CCGCTGAGG AGTNGACTCA GCGCCNANAN AGGCTGAGAA CCTTTEACA
501 CCGATTGCT CCGNCACAT TGTCTCCNC TNGCAGTTG GAAAGAGAA
551 TAAGCCGCG GCGCCCGGT (GAGAGAGAA AAGGCTCCC CCGCCCGGT
601 TTCCCGATA AGGCTCTCC CCGAATCCC TTCTTTTNC CCGGGAAGT
651 TTGGTTTGG GAACCCAAAT TGAATTTCC CCGTCCGTT TTCTTTTNT
701 CCGGCGCT CCGCCGCG (TCCCTGGA AGGCCCCAG CCH
```

SEQ ID. No.: 6

FIGURE 24

30/99

115-96 (3x 0.01)

Session Name: crick.princeton.edu 1

Page 1

```
1  GGGGAAATCT TGAATGOCAT GGGGCTTTC GTAAAGCTTG ATCTCTAANA
11  AAGGGGACCT TTTTITTTTTT GTTTTCTGT TTAATGTTT TTAANAAGAA
101 GGTTCATCAC CACAGAACNC CACACAGAT GAGGAGGAT ANCTTCTGCC
151 GAGCTCTTCA GCTGAGGCTC TAGCAAGAC TTTAGCTTTG TGTGCAAGG
201 GAGGCTTTC CCTGAGGCT GAGGCTTTC TTTCTCTTC CCCAGGCTC
251 GGGGCTTTC GAGTGGCAT GTATGAGCC CAGTATGAG CACTCTCCCT
301 GGTCTTCT CTCTGCTCA TTAGGAGCT AAGGGGCTA NOTTCTTTC
351 GAGGAGAGG GGTCTCTAT GAGGAGTTC CAGTATGAG CTTGGGAGCT
401 GAGGAGAGG TGTCTCTTC TGTCTCTTC AAGGAGCTT AAGTCTGAG
451 TGGGCTTTC GGGGCTTTC GAGGAGTTC GGTGAGTTC AATGAGTAT
501 AAGGAGTAT GGTGAGTAT GAGGAGTTC TTTGAGGAG AAGTCTGAG
551 GGGGAGTAT TAGGATGCT GAGGAGTTC GGTGAGTTC GAGGAGTAT
601 TTTGAGTAT AAGTCTGAG GGTGAGTTC GGTGAGTTC GGTGAGTAT
651 TGTCTCTTC GGGGCTTTC GAGGAGTTC GGTGAGTTC GGTGAGTAT
701 AAGGAGTAT GGTGAGTAT GAGGAGTTC GGTGAGTTC GGTGAGTAT
751 TGTCTCTTC TTTGAGTAT GGTGAGTTC GGTGAGTTC GGTGAGTAT
```

SEQ ID. No.: 7

FIGURE 24

31/99

3949 P01 (50m)

Session Name: crick.princeton.edu 1

Page 1

```
1  ggatccgg: taggaatgaa attagaagct gtagatctca tggagccacg
51  gttaatntg: gtagccacag ttaactgaat tattaccat ctcttgagga
101  tacattttg: tggctgggaa caagagcatg accagcgggt agactgtgag
151  tccctgaa: tatatccgt aggg
```

SEQ ID. No.: 8

FIGURE 24

32/99

scmg

Session Name: crick.princeton.edu 1

Page: 1

```
1  GGGGCTCAAG GAGGCGGATC TTCCGGGTTG AGGGGGAAGC GGGGTGACTG
51  GAGTGGAAAG TTTCGCCAAC MCACTTCTC GGAGGCAACA TATTGGAAGG
101 GACTGGGAGG GGGGGGAGTC CAAATGGAGG TGCGTGAAGG AAAGTTCTCG
151 CCGTCTGAT TCTGAGGCGC GGTGCTGCGC GCGCGCCCTC AATTACCTCA
201 TCGAGCTTTC GTCTTGACAG AACATTCTTC AGAATCCAAA AGGGAAGCA
251 NACTGTTTCG GGATGTTTGA CGGCTATGAT AGCTGCAGTG AGGACACAAG
301 TAGCGCTTCC AGCTCTGAGG AAGTGAAGA AGAATTTGCT CCTTTACCTT
351 CCAATCTCCG AATCATCAAG AATAATGAC AAGTGTACAC ATAGCCAGAT
401 GGTAAATCTC GCATGCTTAC CTGTGAGATG TGTGGGATGG TCGGTGTGCG
451 AAATGCTTTT TACTCTAATA CAAAGCTTT CTGAGTGT TTCTGTTTCA
501 GAANTTACTC GTCCACTCTT AAGAGGACAG CATTCGTGCC NCACTTCBAG
551 TTAGGCTTTT GCGTCCGACH AAGAAAGCCA AATTCCTTCC BAAACBACCH
601 TTAGTTCCTT AATGCTCTGC CTAAGCCCTT NTCCANCTAC CTTCBBAAT
651 CCGCCGAGAC NAAAGCGGCG NATTCGCCA TCTCTGTGGA AAGGTCCTCT
701 GGGGTTAC
```

SEQ ID. No.: 9

FIGURE 24

33/99

24-20 (comiche)

Session Name: crick.princeton.edu 4

Page 1

```
1  TAGGCTGCG AGGAGACGAC AGAAGGATTC GGGCTGGAG AGAGGACAG
51  AAGAGGAGC TTCTCCGCT GGGGGGGGG GGGGGGGAG CTCTGGGG
101 GCGTGGGCT TCACTTTGG GGGCTTCTG TATATGCTG GGTGCTGCT
151 CAGCGGCGC CTCATTTCT TGGCATCTG GAGATGATA NCTTTGATG
201 AGCTGAGAC CGACTACAG AAGCGTATA GAGCATGCA ATACCTGAA
251 GCTCTTCTC TTCCAGAACT ACTCATCAC GGTCTTCTT GTGTCATTT
301 TCTCTTCTG CAGAAATGG TGAAGCTGG CCTCATATG GCGCTTTGG
351 CAGTCC
```

SEQ ID. No.: 10

FIGURE 24

C3-54 (referred in ex. 19)

Session Name: stick.princeton.edu 4

Page 1

```
1  TACGGCTGCI AGAAGACGAC AGAAGGGGAG TGCTGCCCCG TCCAGCGTTT
51  OCCATACCAAT CGAGATGACA TCTGTAGCC GAGAGCCCCA ACACAGCGGA
101  CCCCCGACG: TGTTTTATGA CAGAAATGAA CCCCCGAAAT AGCTTGGCAA
151  CTCACCGATG ATTGACATCC AGACCAATAT GACTGAGCGA GCGCTGAGGC
201  TCTCTCTTTT ACCAGAGGOT CAGCCTTCTT ACCTGTTAGA CATGCGCTGC
251  GGTCTCTGCG: TGAGTGAAA TTATATCTCA GAGAAAGAC ACTACTGGOT
301  GGGCATGAG: ATCAGCGCTG CCATGTTGGA TCCCGCGTTG GACGAGATAC
351  AAAAGGCGA: TGCTGCTAGG GCACTGCGCA AGCTCTCTTT CAAACCGGCC
```

SEQ ID. No.: 11

FIGURE 24

35/99

B4-14 (desig.)

Session Name: crick.princeton.edu 5

Page: 1

```
1  ACCGCTTCGA GAAGACGACA GAAGGCTACG OCTGCGAGAA GACTACAGAA
51  GGTAC:GCT OCCAGAGAC GACAGAAAGG TACGGCTGCG AGAAGACGAC
101 AGAAGGCGCA GGAAGCTGTA AAAGCTATCG TGGGAGGCA GCGCGAGCT
151 TGATTGAGCT TCACTGCTC TGGCAGCCG CTGACCGCGG GTTTCGCGC
201 GGAAGGAGT CAGATATGAA CGACCGGTC GATTATTAG TCACGAGGA
251 AGAGATCAAC CTGACGAGAG GAGCTTCGGG GCTGGGCTTC AACATGCTC
301 GTGGGAGGA TCACGATAT GTCTCGAGC ACGTGGCAT CAGCTCAGC
351 CAGATCAAAH AAGATGGGC TCGGCGGCAH GATGGGCGGC TCACGAGAG
```

SEQ ID. No.: 12

FIGURE 24

36/99

Session Name: crick.princeton.edu 3

A7

Page 1

1 AGCTTTAGTC CCCAGAGATG TGGCACTAC TGTCTCTCT GCTGTGCTG
51 CTCTTGGGGA ATTGAGGCC TTAGGAGGCC AAAGTATCC GTGTCCCTCT
101 TCAATGATAT CACCTTGGAC ACAGATCTT AAAGCCTCT AATGATGCG
151 AACAGCTGG: AGAGCTTCT AGAGCTTCA CCTCTGCTG CAAGCCCTCC
201 TTTTGGCTC TCTCCAGTT CATGACACC CAGTATTTG GAAGTATTG
251 TTTGGGAGG CTTCTCAGA ATTTCACCT TGTCTTTCAC ACCGTTCTT
301 CCACTTCTG GGTTCCTCC AGAGATGTC ATTCTTCAG TTGCGATGC
351 TGTCTTCAAG ATGCTTTAA TCCAGAGCC TCCAGCTCT TCAAGCCCA
401 TGGAGCAAG TTGCGATTC AGTATGCGC CGGCGCTG AGCGCATCC
451 TGAAGCAGA CAATCTGACT ATCGGGGGA TCCAGATGC TTTTGTGCA
501 TTTGGAAGG CTCTGTGGA GCGCAAGCTG ATCTTTGCTT TAAGCCACTT
551 TGAAGGATC CTGCGCTCG GCTTCCCGC TCTGCTGTG GCGGAGTTC
601 AGCTTCTCT GATGCGATG GTGAGCAAG GCTCTCTGA GAAAGCCCTC
651 TTCTCTTTT AGCTCAAGG GATTTCTGA GGTCTGATG GCGGAGAGCT
701 GGTCTCTGG GCTCAGAGC GCGCTACTA GGTAGCTCC CTCACCTTC
751 TACAGTCA CAGCTCTCC TACTGCGAG TCCAGATGA GATGTGAG
801 GTGCGGAGG GCTGAGGCT GTGTGCGAG GCTGCAATG CCATCTAGA
851 CACAGGACA TCCCTCATC CAGAGCTAG TGAAGAGATC CGGCGCTGA
901 ATAGAGCAT TGGGATAT GCTTCTCTA ATGGGAGTA CTTCATTCAG
951 TGTTCAGAG CCGCAAGCT TCCGCTGTC TCTTCCACC TTGTGAGT
1001 CTGTTTAA CACAGGCC AGGATATCT CATCAAGAT CTTCAGAGCG
1051 ATGTTGCTT CTGCTGTTG GCTTCCAGG CTTGATAT CCGCAAGCT
1101 GCGGAGCCC TGTGATCTT TGGGAGCTC TTTTGGGCG CTTATGTGG
1151 TGTCTTGA CCGGCGACA AAAGCTCGG CCGCGCTG GAGTGGCG
1201 GTCTCTACT TGTTCAGA CAGCGGAG AAAGAGGAC TACGAGCG
1251 CAGTCTTAA AAAGAGCCC TGTTAGGCT ACAAGCTAC CCGGCGAG
1301 CAGTATCTT TCTTCCAT TAAAGAACT AAAAAAAAA AAAAAAAAA

SEQ ID. No.: 13

FIGURE 24

37/99

Session Name: trick.princeton.edu 3

Cyt-19 Page 1

1 TCTTGGAAAT CCGGCTGAG GCGAATCTG CACAGGGAA TCTTGAGCT
51 GGAGATCTT AATCATGCT GTTCCCGAG AGCTGATGA GATCCACAAG
101 GACCTTCAGA ACTACTATG GAATGTACTG AAGACATCTG CAGACCTCCA
151 GACTAATGCT TGTGTCAAG CAGGAAAGC GGTCCCGAG TACATCCCGG
201 AAGTCTTCA GAATGTACAC GAGACCTTA GTTCAAGCTA TTATGCTGT
251 GGTCTGACTT TTCTGAGAG GCTGAAAGC TCCGAATTT TCGATCTGGG
301 TATGCTGAGT GCGAGGATT GCTATGTCT TAGCCAGCTG GTTGTGTGAG
351 AAGGACATGT CACCGAATA GACATGACTG AGTCCAGCT CGAAGTGGCT
401 AAACTTAT TTGAACACCA CAGGAAAGC TTGTGTTTC AGGCAAGCAA
451 TGTGACTTTT CTCTAGGCT OCATGAGAA GTTGGCAGAG GCTGGGATCC
501 AGAGTGAGAG CTATGCTATT GTCATATCA ACTGTGTTT CAACCTTTT
551 CCTGATGAA AACAGTCTT CAGGAGCTC TATGAGTGC TGAAGCAAGG
601 CCGGAGCTC TATTTCAAG ACCTCTATG CAGCTTGA A GTCCAGAGG
651 ACATCAATG GCAGAAAGT GTATGCGCG AATGCTGCG AGGCGCTCTG
701 TACTGAGAG ATCTTGCCAT CATTCGCCA AAGATTGCT TGTGCTCTC
751 ACGTTTCTG ACTGCGGATA CTATTACTGT TGAAGACAAG GAGCTCGAG
801 GCGTTCTTG TACTGTGCG TTGTGTCTG CCACATTTG CCTCTTCAA
851 CTCTCTAGA CAGAGCCAGC GAAAGATGC CGAGTTGTT ACAATGAGG
901 AATCAAGGA CATGAAAGG AACTAATTT GATGCAAT TTCACTTCA
951 AGGAGAGGA AGCTTTTCA GTGATGAGG AGAGGCGAG TGTCTGAGG
1001 AACTCACTT TGTCTCGGA TTGTCTTTC ACAGCTTTG AGGCTCTCT
1051 GCGAGCTCC CAGGCGCTT TGTGAGTAG AGACAAAGT TCTAATCAGA
1101 GATCAATCA AGCTTGAGA GACTCTGAC AAGATGAAG CCAGACATGC
1151 AGCTGAAGC AGCGAGCTT GCTGTGGCA GAGGAAAGC TCTAGATCT
1201 AGAGCTAGC CCGAGCCAC GCGGCTGAG AGGCTGCTA AAGGACAGT
1251 ACAGAGCTT GTTACCTGC TTCTGCGAG TCCACAGAT ATGTGAAGT
1301 GCGAAAGCA CCACAAGCTA GACCACTCT AAGATAAGA GTGACTTTA
1351 GAGATTTTA ATGAAGGTT CACAGCAAT CCGCTGCTT TCTATTCTC
1401 TATTTCAAG TTCTGTGCG AGCTAGTGT CAGAAAGTA ACTTGAAGC
1451 CCAAGTTTA CTCAAAGGC AAGAGCATC ATCAAGCTG TGAAGATTAT
1501 CTCTCTCTG GCTACGACA GAGACCTCT GGTCTCTCT CTGTGTTAC
1551 CAGGAGCTC AATCTTACT AATTTATG TAACCATAC AAAAGATTOT

38/99

Session Name: crick.princeton.edu 3

Cyt-19 page 2

1601 CAGCTCAAT TTCTAGAG TATTCTTAG GTTCTCTCT GCAATTCTT
1631 TCTCTCTAC TGAATGAAA AGAAGACAA TAAAAATAA ATTGACTTC
1701 GAAATGAAA AA

SEQ ID. No: 14 cont'd

FIGURE 24

39/99

Session Name: crick.princeton.edu 1

SA 61

Page 1

1 CCGAGCCCTG GCGGAGTCTG TCGAGCCCCC GGTCAAGAC GCCATCTCT
51 ACCAGCAGCT CGTAAAGTTT GCGAGAAAT GCTGCGCAA AGTGTGGCT
101 CTCTGTATG CCGGAGCCCC ATCAGGCTA GCTGCGTAG AAGCTGGA
151 CGTGCTGA1 GGTGCGCTG GACCAAGAG CGACAGGTC ACAAGGCGCA
201 CCGTCTGAG GAGAGCCCCG GTCATAGCT TGGTGAAGT TGTATCTTC
251 CTGCTCCG ATGCGGAGAG CTGTCCAG GACAGTGT CTTCTCTAT
301 TACCACTAC GAGCGAGCC ACCTGTTCG TCGACAGAC CCGAGTCT
351 GGTCTGCT CATCTCTAG CTGCGTTC CCGTACCG AGATCTTCG
401 TCAGATCA GACAGCTGA GATCCAAA AGGCGTTT TTCCATGA
451 AGAAATCT ATCTACTCT CTTGCGAG AGTACCGAA TTCTCTGA
501 TGTGCGAA GACAGAGCC ACCTGCGT CCGAGTGA AGAGCTTC
551 CTCTCTTC TGGCGAGA TGACATCAA CTGAGGAG CATCCAGCC
601 CCGAGCTCT TTAGCTGC CTTACGTT CTTGCGAG TACGCTCT
651 ACAAGGCT GTTCTCTTT GAGCTGCG CCGCTCTGA CTCAAGTGA
701 GCGCTTTT CTTCACTAG CCGCGTCC CAGACATAT GTGGGTTGT
751 GGTGCGCG ATGCGCGC AGCGGAGG TCTTCAGG CTGCGCAT
801 CCGTACTCT CCGCTGCT CCGGCTCT CCGTGGCT CTTAGAGCC
851 CTTGAGAG TTGCGAGG CCGCGAGG TTTGAGTC CACTGCGG
901 AAGCTTCT CTAAGTAT CCGGCTCA AGCGTACA TTCTGCTCA
951 GCGCTACAC AGAGAGAG GCGATCGGT CTCTATCGT CCGTGTGAG
1001 GCAGAGAG AGGACAGAG GCGCGCGA GCAATTTAT GAGAGCTGT
1051 GATGCTGA GCGAGCTT GCGTGAAG ATGCGGCTC TGAAGCCAA
1101 GAGGCGCT TTGTGCGG CAGCGCTT GCGAGCTT ATCTACATA
1151 AACTGAGA TCTGAGTTC CCGGCTCG CCGAGAGG CAATCTGAA
1201 GCGTGAAG GAGGCTCT GGAAGTGA CTGATGAG CCGAGGCG
1251 GCGGCTCT GAGGCGAG GAGGATCA GCGAGCTG GTGAGCTG
1301 TGAGCTTA AGGAGAGG GCGCGCGC CTTGTGAG CCGTGAAG
1351 CTTGAGTC CAGGAGTTC AGGAGAGG TGATACCCA AGAGGAGAA
1401 CAGTTTGA CATAGCGT TACTCTGAG CTGAGGAG AGGAGGTT
1451 GCGGCGAG GAGGAGCT TCGGCTAG TCGGCTCA GACTGTAG
1501 ATTGAGAT AATAAGCT TCTATCAC TTCAAAAA AAAAAA

SEQ ID. No. 15

FIGURE 24

Session Name: crick.princeton.edu 1

Cyt-28

Page 1

1 AGAGACAGG TATCCCTGC CTCCACCGG GCACTTTTA CTCTCTAGG
 51 AAGAACTCT CAACTCCAT GAACTCTGA GACTCTGCA AGCCTCCTA
 101 GGAACCTA GCTTTTGA (GCACTTGG CACCGAAGC TACCCCGTA
 151 GCGGAGCTT CACTGGATT GACTTCAG CTCCCTAGC AGCCTCAGC
 201 TAGCTCTG GCACTTGG GAGGAGAG ACATTAGAG CAAAGCTGT
 251 TAACAGAGT CCTCCCTAG CTCTGAGG AAGAGCTGC TTTTCACT
 301 GCGGAGAG GCTGAGAT CAATGCTGT CCAAGTCTG CCGCAGATG
 351 TCTACTTCT ACTGATCTG TTTTCTCTG TCAAGCTGC ACACAGTGC
 401 AGCCCTGAG AAGCTTCTG CTCTCTCTG CAGCTGAGC AGCTCAGC
 451 GAGCAGCTC CACTGATG ATTTTCAG AGCTCAGC TTTCTCTGA
 501 AGACAGAG GACCTCAG ATTTGCTGC CTCTCTGAG AGCCCTAGT
 551 ATTCCCTCT TCTCCAGG GCTTACAG CTCTTACT TCTCTCTA
 601 CTGAGTCTG CACTTGGG GACTCCACT GCGCTATGC AAGCATGCT
 651 AGCTCTCTG TACCAAGCT TCCAGCTGC TCTCTCTCA GAAACAGAG
 701 CAGAGCTGA AGCAGAGC GCGCTGAT GCACTCTG TCACTCTG
 751 GCAATCTCC CAGACACTA AGCTCTCTG GCTCCAGC TTTCTCTT
 801 CTCTCTCA TCCCTCAG AGCTCTCTG ACAAATGAT TGTGAGATG
 851 TGTATCTCA AGAAGGAT GCAAGCTT AGCAGTAC TCAAGCAGC
 901 TCAAAAGT GCAAGCTG AGCTCTCTG GCTCTCTG AGACAGTCA
 951 AGAGCTCA GCAAGCTG AGCTCTCTG GCTCTCTG AGACAGTCA
 1001 TCTTTTAG AGAGCTCT CAATGCTCA GTCTGAGAG TCCACCTAG
 1051 AGCTCTCTA GAGATCTG ATATCTCTG CCAAGAGAG GAGGAGCAG
 1101 GTAGCTCTA GCACTCTG CTCTCTCTG CCGGCTCTG ATCTCAGAG
 1151 AGCTGAGCT GCTCCCTGA TACCTCAG AGCTCTCTG TACTAGCTT
 1201 CAGTCTCTA GCTTTCTG AGACAGAG TTTAGCTA GCTCTCTG
 1251 AGAGCTCTT GCTATCTG GTGAGAGC CCAAGTCA CAGCTCTCA
 1301 GATCTCTG TACTCTCT TACCTCAG CCTCAGCTA AAAATCTG
 1351 TCTCTCTG GTCTCTCT TTAAGAGCT GCACTCAG AGCAGCTCA
 1401 GCTCTCTG TCACTCTG TACCTCTA GCACTCAG ACAGCTCT
 1451 TCTCTCTA AGCTCTCA CACTCTCT GCTCTCTG TCTCTCTA
 1501 AGAGCTCA GCACTCAG AGCTCTCT CAGCTCTG TCTCTCTG
 1551 GCTCTCTA CTCTCTCT GCTCTCTT TACTCTCT TCTCTCTG

SEQ ID. No.: 16

FIGURE 24

41/99

Session Name: crick.princeton.edu 1

Cyt-28 Page 2

1691 TCTTCCAAGA GGAAGTCAGC TAACTACACC ATCAAAGTCC ACATGAACCT
 1691 GCTGTCCGCT GTCTTCTCTC TGAAGCTGAG CTTCCTGCTC AGCAGGCTG
 1701 TGGACTTAC: GGGCTCCAA GAGCCTGTC GCACCACTCC CATGTTCTG
 1731 CACTTCTCC: TCTTCCCTG CTTCTCTCTG ATGGGCTCG AGGCTACAA
 1801 TCTCTACCA: CTGTGCTGAG AAGTCTTCCG TACCTATCTG CCGGCTATC
 1851 TCTTAAAGC: GAGCATCTCT GCTGGGCTT TCTCTCTCTT CTTGTTCACT
 1901 CTGTGCTG: TGTGCTATCT GAATAACTAC GCGCCCATTA TCTTACCTCT
 1951 GCGGCTGAG: CCGGAGCTG TCACTTACCC CTCTATCTCT TCGATCCGCG
 2001 ACTGCTCTCT: GAGCTATCTC AGCAAGCTCG GCTCTTCTAG TCTGCTCTTC
 2051 CTGTTTAACT: TGGCTATCTT GCGGACCATG GTGGTGCAGA TCTTGGGCT
 2101 TCGCCCTCA: AGCGGAACT GCGCCGAGCT GCTGACCTCT CTGGGCTCA
 2151 GCTGCTCTCT: TGGCTCTCCC TGGGCTTGG TCTTCTTTC CTTCTCTCT
 2201 GCGGCTCT: AGCTTCTCAT CCTCTACCTC TTGAGCTCA TAACTCTCTA
 2251 GCAAGGCTT: CTGATCTTCT TGTGCTACTG GTTCATGCGG TTCCAGGCCC
 2301 AAGCGGCT: CTGCTCTCTG AAGAACACT CAGACAGGCG CAACTCTCCC
 2351 ATCTGCTCT: GAGGACCTC CTGAGGCGG ATCTAAGCCA GCGGACAGC
 2401 TCTCTCTCT: GAGGACACT CCAAGGCTC GCTCAGGAT GTCTGTGCGC
 2451 CAGTCTCT: CCGAGGCTC CTCTTCTCTG TATGGGATA CTAGAGAGG
 2501 GCTGCTCT: TGAAGGCTA GGGCTCTCTC TCTGTTAGGT AGATACCTAG
 2551 GTTCTCTCT: GAGGACTCT GCTCTCTCAA GGGCTCAGAA GCACACTGCG
 2601 ATTCTCTCT: TGGGCTCTT TCACTCTCTA GCTAAGGCTT TCTCTCTCT
 2651 GCGGCTCT: GCTCAGCTG TTCTCTCTCT GTTCTGAGAC CTGAGAGCC
 2701 AAGCTCTCT: TAACTCTCTA AGGAGGCTGA CACATCCGCG CTGAGACACA
 2751 GCTTACTCT: TTGACTTCTT GCTCTCTCTC TGTGCTCACC ATGAGATCC
 2801 GCAAGGCTC: ACTGGGCTA AATGTTCTCT GAGAGGCTTT GAGGCGAGG
 2851 CAGTTTACCA: GCTGAGCATC TCGGCGGCG TTCTCTCAA CCGTCTCTTT
 2901 CATCTCTCT: CCGGAGCTG TCTCTCTCTT TCGGCTAAG CTGAGCTGA
 2951 AGCTGCTCT: CTAAGACCA TCTCTCTCTT TGGGCTGTA GTTCCAGCA
 3001 GTTCTCTCT: GCGGCTATC AACTCTCTC TGTCTCTCTG GCTTCTGCT
 3051 CTGACTCT: GAGGCTTCT GTCTGAGGCG TCTCTCTAG CTGCTCAGC
 3101 TTTCTCTCT: CCGAGAGG AGCTCATCT CTCTGAGG CTCTCTCTC

SEQ ID. No: 16 (cont'd)

FIGURE 24

42/99

Session Name: crick.princeton.edu 1

Cyt-28 Page 3

3151 TGGCACTAC: TGGATACAG CCACCTTTC AACCCAGCAG TCTGAAGACC
3201 AAGACAGCC: CCTCTGCTGA CACTGGCCAA GCTTGATCTT TTTCCTAAGA
3251 AGTGTCTT: AGATCCCCC AGGTGCTCA GAAGACACTG GGTGCTTAG
3301 TGTGAATTC: GTCTACTAA COTACAGTA GCAGCTCTC ACCCCCACCC
3351 CCCCAGAG: TCTACCAAG TCCTGAGTG TCAGGCAGG GGTGGAAT
3401 CCAGGAGAG: TTCTGCAA AGGCAGATT TCATCTTGAC CTCAGCTTC
3451 AGGTTCGG: GAATGTTCTT TTTAAATACC AGTTCATTG TCTTTTGATA
3501 TTAAGCTCT: TTATAGAGAG TCTGAACT GTAGCGATT GTGAGAAGA
3551 GAAATAAA: TGACTCTTA TCTATGCA TGGCAAGCA GCACAAAA

SEQ ID. No: 16 (cont'd)

FIGURE 24

43/99

AFTP2

BMP1 TGF β /PDA

Seq Length: 430 February 24, 1997 11:27 Type: N Check: 1785 ..

1 ACGGACATGG CCGCGCCCGC GCTCCAGGCC CAGCCTECAN GCGGCTCTCA
31 ACTGAAGTTC CTACTGTTCC TGCTGCTGTT OCTGCTGCTG CTGTCATGGC
101 CATGCCANGG GCGCNCCTG GCAATGCTG AACAGCCACC CTCGCGCCCT
131 GAGTCCCAAC TCACGCCGA CAACTACCG GTTCNCTTC AGGACCTGCT
201 GAGCCCGCTG CATGCCANNC ATAGCCGAGA GGACTCTAAC TCTGAACCAA
251 GTCTGACCA TTGTCNCGA TACTCANTCC ACANGTGAGA TTGCGGTCCA
301 CGGCCAGCTG CTACTCNCG TCAAGCGGCG GTGCTGANT CANGTCTCC
351 CCGAACCTAC CAGGTGCANC NAGCGCTGCT GTGCTGACH NCGAAGCCG

SEQ ID. No.: 17

FIGURE 24

44/99

D77 - T404/PE4

1 CGGACGCTTG CGGAAAGCCG GTGGGATAT AGGTGTCTG GTGGCTACA
51 ATTGACAGG CCTTTGGGA GGGTGACCA CTTGCACTG TACTGTACT
101 TCTAATGAG AATGACTAT CACTAAATA ACCTGATGA AGAAGATTG
151 AGGTGAGCC CAGCTCTTG TGGTGTCTT CCACTCCAG AAGGGCCCA
201 ACATCAAGA GCGAGAGAG GTGAAATCT TGGTGGCCA ACAAGATCTG
251 AGGAAGCAT CTCTGGCAT CTGAACTTA AGTGTAAAG AGAAGCAT
301 CTATTAATG CAGATGCA CATTCCGAG AAGCAATGA AGCACAATG
351 CCTGGCTGA AGTGAAGCC CGACCTTAA AACTGCANA AGCCTGAC

SEQ ID. No.: 18

FIGURE 24

45/99

1 CJAATTCCCG GGCTGACCCA CCGGCTCGAG AGGCTAGACA GAAAGCCCGG
 51 AAAGGCTCTT GAGTCAAT OGACCCCATG GAAGTGAACA ACCTCAGCAT
 101 CGAACCCGAC GAGGACAGCT GCAGCCGCGA CAGTATTGAG GACAGCTACA
 151 CCGCCATGCA AAAGTCCGAC AAGGACGCCA TGAACATCCA ATTTGCTAAT
 201 GAAATGCGCG AAAGTCAGAA GTTCTGACA AATGCTTTT TAGGGAAGAA
 251 GAAGCTAGCC GATTACGCG A. TGAGCATC ACCCTGGAAT GACTTCCTTT
 301 GGAATGCTCT CATTTAAGCT GAGCTACGCC ATCATGCGCA GTGGATCTTT
 351 AGGCTTGTCC TAGCCATGG CCAACACCGG GATCATCTTT TTTATAATCA
 401 TGTGCTTAC TGAGCAATA CTCTGCTCT ACTCGGTCA CTTTTGCTG
 451 AGACAGCCA AGGAGGAGG CTCTCTAATC TATGAAAT TAGGCGAGAA
 501 AGCATTTGGA TGCGCTGGA AATTTGAGG CTTCATCTT ATTACATGC
 551 AGAATTTGAG AGCCATGCA AGCTAGCTCT TCATCATTA GTAGGAGCTG
 601 CTGGAAGTAA TCAGAGCTT CATGGAGCTT GAGGAAACA CTGGGAAATG
 651 GTACCTCAAC GCAACTACC TGTCTTATT TGTGCTGCGT GGAATCATC
 701 TCCGCTCTC TCTCTTAAA AATTTAGCT ACCTTGGCTA CACGAGTGA
 751 TTCTCTCTCT CCTCATGCT GTTTTCTCT AGTGTGCTA TTTACAAAA
 801 ATTCCAAATY CCTGCTCTC TGGCTCTCT GATCACAAC AACGGAAATC
 851 TGAGCTTCAA CACACACTT CCGATTCACA TGATCTGCTT GCTAATGAC
 901 TCGAGAGTCT CCGCTGTGAA CTTCATGATG GATTAGCTC ACCACAACCC
 951 AGCTGGGTG GATGAGAGG AGGTGCGAG CCTCTTCAC AGCAATGGCG
 1001 TGAATACGA AAGCCAGGT GCTGAGAAAT GCAACCAAA ATACTTTGTG
 1051 TTCAATTTCC GAGCGGCTA TCAATCCCA ATCTGCTT TTGCTTTGT
 1101 CTGCCACCTT GAGGCTCTT CCATCTACAG CGAGCTTAA GATCGATCCC
 1151 GCAGAAAGAT GAGAGCTG TGCAACATTT CCATCTCAG CATGCTGCTC
 1201 ATGTACCTT: TCGGCGCTT CTTGGTTAT CTGAGCTTCT ACGGGAGCT
 1251 TGAAGACGA: CCGCTGATG CTTACAGCA GGTCTACACA TTTGATACGG
 1301 CTCTCTCA: GCTGCGCTG GCACTGCTG TGGAGTGA ACATGAGCG
 1351 AGAGCCCAT: GCTGCTTCC CAGTCTGCT ACTTCGTTG TATCAGCTA
 1401 GGTGTTTTC: AAGGAAACCC TTTGAGGCG TGAAGCATTT CAGGATCTC
 1451 GGTGAAATCA TCATGCACT CAACACATC CTGCTATCC TGTGCTTAC
 1501 CATCAATTA: ATTTTGGAT TCATAGGCG TCTTCTGCC ACTATGCTA
 1551 TTTCAATCT: TCGGCTGCG TTTATCTCA AGCTGCTCA GAAAGACCT
 1601 TTAAGATCA: CCAAGAGAT TGGGCTTTG GTCTCTCTT TGAAGCAAT

SEQ ID. No: 19

FIGURE 24

46/99

AFTP1

1651 TATTTTCAG ATGGGAAGCA TGGCGTCAT TATACTCGAC TGGATCTACA
1701 ACCCGCCGAA TCCTAATCAC CACTAATCCC GGGGAGAGCG GTCTCCACTG
1751 GAAACAGCTG AATTTGTCTG AAGGACATTT TAGTTGTCTT GATTOGATG
1801 TTAGTCTGAG GATTEAGCAA GATTCCAAAG ACCTTTTCTT AGCTCTATCA
1851 GCACACATTT TACCCCAGCG COTOCAGTGC AGTGTGTGAT GCCCGAGTTO
1901 TGTTCGAGC AGTTGTGCAA GCTGAAGCCT GTTGGCTGCG TGTGTTGGTC
1951 AGCAGACAAT AGCTGTCTCC CCCATGCTCA CTCGACTTCT CTCGACCCCG
2001 AGATTAACAG GTGATTTCTAC TCTCGAACA TCAGACAAAG ACCTCTGCT
2051 TGGGATCTT GTGGAGAGA AATTATGCG TTTGTTGGG AATGTTTTG
2101 TTGGGAATGG TGAAGGATGC ATTAAJAATT CTGTGGCACA CATTTTACC
2151 CAGGCGCTGC AGTGCAGTGT TTGATGCCC AGTTGTGTTT GCAGCAGCTG
2201 TCGAAGCTGA AGCTGTGCG CTCGTGTGT TGTGAGCAG ACAATAGCCT
2251 GTCCCGCAT GGTCACTCCA CTTCTCTCCA CCCCAGATT AACAGGTAAT
2301 TCTACTCTCA GACATCAGA CAAGACCTC CTCGTTGGCG AAGTATCATC
2351 AATTACGCG ATCTCTCACT CTACACCAAC ACTAAGGCTC GGTGACTAG
2401 CTGAGCGAGG CGGATATCTT GGGCTGTCCC TGTGAGGATC ATGACGTATG
2451 ACGGTTGCA GTATAGAGTA CTTCAATTCA ATACTCAAGG AATAGTTTGC
2501 CCAACCTACT TATTACACCG AGTTAGTGA

SEQ ID. No: 19 (cont'd)

FIGURE 24

- GAGATATCT GTTGGAAATT GTGGGTGGA GTCAGCTTC GAGCGCTGC
 51 GAGATATCT GAGAGAGG GCTTCAGCA GAGTACGAC TGGCGCTGC
 101 AGCGAGAG GAGAGAGG GAGGAGCTT GCGCTCTAG CTTGATCCCT
 151 GTCTTGCTG CTTGAGATT GCGAGCTGA GAGATGCGT TCTGAGAGG
 201 TGGATGTTA CTTGGCAGA GAGCATTTT GAAAGCTTG AAGAGAAAT
 251 GGTTCGACT GTGTCAGAC GTTCAGCTA TCTACTAGA TATCAGACT
 301 CCGCAGAGC TACAGAGAA GTTCAGATG CCGTGGACT GATCAATAT
 351 CCGCAGAGC CATGAGTGC GAGCATCCA GCGTCCAGT GCGAGGCCA
 401 GAGATGCTT GCTGAGATC GTTGGCGAG AGCGAGAGC CATGAGTCT
 451 TGTGAGAGA GACAGAGCA TTGCTGCA TGGAGTTTA CACTCAGGA
 501 TTGAGAGC AACAGACTT AGTTGCTTC AGCAATCTG TCTGAGAGA
 551 CTGAGTGC CCGTCCCGG GTCGCTAGG CAACCTATG TACAGGACC
 601 CCGAGCTCT AGGCTATGG TCCATACAG GCGCGTACC CCGAGGAGC
 651 TCAAGTTCT TATGCGGCA AGCGAGAGC ATATCAGTG CCAAGCAGT
 701 AGCGTATGC AGGATTTAT GAGAGAGC CTGCAACCA AGTCATCAT
 751 CCGAGCGCT AGGAGAGCA TACAGTGC CTGCTCTGG GATGCTGCG
 801 CCGGAGAGC AGCGCATGG CCGTGGCTC TCTGTTCTG GTCTCTAGA
 851 GCTTCAGC TTTCTGTC ATAGCTCTG TTAGTCTGT GTGAGTAAT
 901 TTGATTTCA GCGATTTCT GTTGTGACA AGTGTCTTC ATAATTAAT
 951 AATAGTCTT TTTGAGCTG GTAATCTAAT AATTGTGACT GAGCTGCATG
 1001 GTACCAACA GAAAGCCCA GGTATGCTG GATGAGAGC CTGAGTCTT
 1051 CCGGTTCTA GCTTCAGCA AGTCTTTCT AGGACTTTT GATGCTTT
 1101 ATGTAAAC: AGCAGTTAA ATGCGCAAT TCGTCCAGT TAAATCAGT
 1151 GTTGAATTA GAGATGCTT TCGTCTAT GCAATACTA ATACTCTGA
 1201 TGGAGAGC TGTGTGAGG TGTGTGAGG AGAGTACAG CTTCTTAAC
 1251 TGTGATTTT CTTTAGAGC CCGTCTGCT GTTACCTAG GAGCTGTGG
 1301 CTGCTGCTT CTGAGACT ATGCTGTAG GAGCTGTAA CGTACCTTT
 1351 GAGCAGTTA GGTACCGCTT GAGCTCTA GGTATCACA GCAAGATTG
 1401 CTGCTGCTA TCGAGAGG CAGCGCTCT CTTTAAAT TACGCTCAG
 1451 TAACTCTC: AGTTTAAT TTTGTTAAT CTTCTGTTT CTTTCTG
 1501 GAGTACTG CATGAGTCT GAGTTAGAA TTGATGAG TCGCGTGG
 1551 TGGAGATG CAGGAGCTG TCGAGGAG TGTCTCAT GTTGTTTA

SEQ ID. No: 20

FIGURE 24

48/99

1601 GGTATATTGA TTGCTAGCTC AAGCTGCAGA GAGCCTATAG AGACTATTTT
 1651 TCTACTTOTT AAGAAAGTAT GTTCAGCGGA ATTGAGGAGA GCCTTGTTTG
 1701 AATGTTCTCT CCTCAGGCGT CTTGCGCGCC AACTGCGTG GTCTGCGGA
 1751 GCCTTTTCTT CAGCAGTGA AAGCAGCGCC CTTGCGTGT CAGAGTCTCA
 1801 GGTTTTGAGT GAGAAATGG ATTGCGTAGA GCATCTCAGG GATTGTTCTT
 1851 AATGCTCTCA GTTATGCGGA TCCAGCGTGT TTCTAGTCC AGACCGCTC
 1901 ACCTCAGAA AGCTTAAAC ATTCTGCTC CCCAATGTG TCGACTCTG
 1951 AGAAGCTCA AATGCGCTT TCTAAGGAA ATTGTATTT CTAAATTAG
 2001 AGAATCATT TTCCAGCTAT TTAAATTTA TGTCTTTCA TTTTAAAG
 2051 CTCCCAAT: CAGCTTTGT GTCCCATAT TTTAGTTCA AACAGTACC
 2101 CCGCAGAA GTGGATGCG ACAGTGAAG CCGCGCGCC TCTACTTTG
 2151 ATAGCTCTT TTCTCCAGG TGCTCCAGG AAGCATTCAT TCTGACTTG
 2201 CTCAGCGCA TCGATCGTG CTGCTTGGC GCGCTGCTG TGGTAGCTC
 2251 TTCTTTGTC AGATCAAGT TTCAACAGT CTCCATGTA CAGAGTGGC
 2301 AATAGATGA GTTGTGCGC ATAGCTTTT CTGATATC CTGCGGACC
 2351 TTTTGACAG TCGCATTTT CAGCTGACAT TTTTCTTCT GTCATCTCT
 2401 ATAGATGGA TATGTGACA CATGCTACG ACCTGCTCA GTGCTGCTT
 2451 AATAAGCAT ATGCTGATT CACATCTGT GCTGTATGAC TCGCATTTG
 2501 TCACAGTGT ACCATTTCT AAGCTCGTG CTTACTTAC AAACACTAA
 2551 AAGCAGTGT TACTTTTCA CAGTGATTT AATTTTACG TTAGTACTG
 2601 GCAATGTTA AGCATAGAG TACTGAGTCA CATCCCTGA GTACTTTTA
 2651 AACAGATTT TCTGCTACT TCCATGGT GTGCGCTGC TGTCTCTG
 2701 CCGCAATGG GCTAGCTTA CAGGCAAGC ATAGTTGAG CTGATCATC
 2751 GTCTACAG TTTGACTTA TTATATACC AGAATGAAT ACATTTCTG
 2801 GCACTTCA TCTCAGCGC TGATCTCAT AGACCGCAC CTTCGATGG
 2851 CTTTCTGCG GTCACTGTA CTTCACTGAG TCTCGCAT CTTCAGTTAG
 2901 AAGGAGGAG TTTCCAGCA TTAGGAGCA TGTCTCTCC ACTGAGACTT
 2951 AATGATGAT CAGCGCTGG AAGAGCTGG CTCTGACAC TCGATCTG
 3001 CTGATGCTT TGTCTCTTA TTCTTTGAT TTAAGAACCT TTCATATGA
 3051 AAGCTGAG CTGCTCAGA TGTCTCTG CCAAGAGGC CTGCTTAGG
 3101 TCAATGATC CACAGTAC CTCTGAGT TAGCAAGT CTTGCTTTG

SEQ ID. No: 20 (cont'd)

FIGURE 24

49/99

DDI 36

Session Name: url:k.princeton.edu 1

Page 3

315: AGACAGATG GTTCAGATT ATTTCCTACA TCTGTTGTTG ACCCATOCA
320: CCTCTCAT: TTGCTTCCA GTTACOTAG ATGAAGATG AAGGCAGAG
325: GATCAGAC: GTCTCTTTG TATTTCTTC TTTTATCTG TTGCATCTAC
330: CCGGCGCTT: TTCTGCTGT CTOTCATAC AGTATGTTA TAACTGAAT
335: TGTAAATA: TAAATGATC ACTAACCTA AAAAAAAAAA AAAAAAAAAA
340: AAAAAAA

SEQ ID. No: 20 (cont'd)

FIGURE 24

50/99

1 TACCGGTCCG GAATGCCCGG CTGGACCCAC GCGTCCCGGG CTCCGAGAGG
 51 ACCGACATTA CGAGGCCAGG CCGGCCAGGA CAGCCACTGC AGGCACTGAG
 101 AGCCTTTCAG TCAGAAAGGA AGCAGAGACA GGAAGAGGCA GTTAGGAGAA
 151 AAGGAGCAAA ACTGAGACAA CCAATGGAGA TCCCTGAGGA GTCTGCGCA
 201 AGCCTTTCAG TGTGCTTAAA AGACATCACT GAATTAATGG AGAGACATCT
 251 CAGTCAGATG GAAGCGACCC TGTCTCAGAG TCAAAAGCTC TCAGATGGAG
 301 AGCTGTGAGG ATGGGCATCT GAGAGGCTGG TCTGCGAGGG AATTATAGG
 351 ACCAGGAGCC CAGAGAGGCT GATTGAGAGG AGAGAGAGGC TACTCAGTGT
 401 CCGTATTCAG TCTGACTTGG CTGGCCGAGT AGTGGGCA CAAGTAAAGG
 451 CAAAGGATTT TTCACTTTTT CAGAGACAGG CCAATTTTAC ACAGACTATA
 501 GAGAGGTTGG GTTTCAGTTC AAGCCCCCTA GTTAAAGGAG AAGGTTGGGG
 551 ACTTAGTCTA CAGCTGCTA TGGGTCACAA CAGACAGACA GAATCTGAGG
 601 ATAGCTACCA GTCCCATTC AAGCAAACTT ATTTTTCTC AGCCAGGTTT
 651 AGCTACATCC CATTCGCCAC CTGCCATTTT CACATCAATG ATCTGCACT
 701 CTCCAGGCTT GCTCTCCAGG AACTAAAAAG TATTGAAGAA ATCTGAGGC
 751 AGACTACACA CCACCGAGAT GCACTACCTT TACTGAGGCA CAGGCTTAAA
 801 AACTTTTTC CAGAGTTTGG GTCTCATGCT AACCAGGCC CTGTGCACCT
 851 GCGGAGATC TACTGCTGGA AAGCCATTC AGAAGGTTT AAAAGTGAAC
 901 ACTTGGTGA TGTAAAGCAG CAGCAGAGG AGTCTTTGAA TATTTACATT
 951 ATGAGCAGIT ATAGTGGCTT TGAATTTAAA GTTGTGCGA GTTAAATAT
 1001 AACAAATCA AATCAAAAA CAGCATTTTA CAGTAAACT CATCTAACT
 1051 TCGAAATCA GGTACAACTA TCTGTAGCCA AGATAAGTGG ACCAGCAGAA
 1101 GCAATAGCA TTTCCAGTGG CAGACCTGCG CTTGTAGCTA GCAATCAAG
 1151 CTGATCTGTT ATTGATAAGG AACTGCAATT GGTACCTATT TGGGACATTA
 1201 TCTGTCTGAG TCAGAGAACT GAATTTAAGA ATGCTCTTCA ACTGGCTAAC
 1251 TGGTCAAGG ACCACTAGAC TGTCTGACT GAAGTGGTGG CCCAGATTC
 1301 AGAAGGCGA GAATTTCTGA CTGCTAGAAA AGAAGCTAAG CTTTTCCTAA
 1351 AGATGTGCA AGGCTGGGAG GTTCTGATC CTGAGAGACA GCTTAAAGAG
 1401 TTATAGAGAT TTATGCAAGC ATTGAGTCAA AAAATAAAA GTTATGACAT
 1451 TTGATTAAGC ACATGCTCA GAGATTGGA TCTOCAAAAT TTTCTAATA
 1501 ACAATGCTCA CTCTGCAAA AATTCACCCA GTTATAAAGC TCACTTTATT

SEQ ID. No: 21

FIGURE 24

51/99

Session Name: crick.princeton.edu 3


Page 2

1551 AAATCTTACT TGTGACGCT TGTGACGCT CATGCTACA AAGTGACAAA
1601 CTTTCTCTCA GCACAAATCA TCATACAGTG GATCAATCAG TCAGAGTCAG
1651 AAGGAGTACT AGTCAAAAT AGCTCATTTT CTGATTTTAT TAACACCTTA
1701 AAGAAACCG ACAAATACCT AATGGAAGAG AGTTTCAAA CTGAGCCCCC
1751 AAGAAAGCTG GAGAAACCA AGAGATGCG TACATATGAA GTACCCACAG
1801 CTCTCACTC GTTCTTTGAG TACCTCAGAG AAACACAGCA GTCAGACATG
1851 CAGCTGCTG TACTCTCCAT TGCTACTGCT GTAGGCTATC AGTTGCTTAA
1901 CAGTATTTT CAGCATCTT TGGGCTGTA TGAGTTTAACT TTCTCTTGG
1951 ATCAATGGA AAGTACCTA CATTAATTC AAGTACTGA AATATTTTC
2001 AATTACGAG CCCAGGCTT GTTGGTCTC ACAGGCTTAA GAGCCACAGT
2051 TGAAATCACA CATGTTTCTA CAGAGAGAA AGGACACCT TTGACATTAA
2101 TACAGCTACA TATGGGCTCA CTCTCTCTG AAGAGTTTC ACATGCTCTC
2151 ACATAACATG GAGAACATCA TCACTGGCA AGGCTGGAGA ATGATTTGAG
2201 ATTCTCTAT GAGGGGACT ATAAAGCCAC CAGCCATTC TTACAAATGG
2251 ATGAGTAAA AAACAATTG GAAATTTTAT GGCATGAAA GAAACAGACT
2301 TATTAATCAC AAGTAATGA AAACAGACA AAGAAATGA TAGAAATGG
2351 ACAATTCTG GACTTACTTC AACGTTTAC GCTAGACAT TACTATCCAA
2401 AAAAAAAAA AAAAAAAAA AAAAAAA

SEQ ID. No: 21 (cont'd)

FIGURE 24

52/99

S. abt/LL2-

Session Name: crick.princeton.edu 1

Page 1

1 ACTAGACTCT GGTAGAGAAA AAGTGCCTC GAGTGGCTG CTGACACCC
51 TCATGACTCA CCAGATOTO GAGCTTTTCA KAAAACTCA CTTFEAAAGC
101 TCAGCAACA GCTCATTOGT TCTCTAACT TATTGTGAT GCAAGTGGAA
151 ATAGATGENT ATACAGCTCT TAAAAATOG ATGTTCCCTC AGCTGCTGCC
201 NTGCTGGAA' GGGGTCTTTG AAAACAGCTT TTGACAAAA CAGATGTCTG
251 GTTTTCCAAG: TGAAGAGAG AATTGAAAG GAGGACTTC CNTGAAACTG
301 AACAGGGAJ, ACCAGTGGC GCGGTGTC GACATTTA GCTACAGTA
351 CATTATCACT' GATCTGGCTT CTGCAAGAT CATTGAGAG GATTCCTGG
401 TACCTTCAG, ATGCTGGCG GAGTGTATA AACAGAGTG GCTGCTATG
451 CTAGGGCTG AACAGAGAG TGAAGTGGC GCTCAAGAA TCATTAAGA
501 AGAATTGAK: GGAACAGCA TGAAGTGTG TGAAGCTT GCCAAGATG
551 GTGATTACT: CTGGGCTGG ACAGCTTCA ATTGCGCTT TGAGCTCCTT
601 GTGACTTAC, CCAATGATA CATCATTTT CAAGCCATA CACTGAACCA
651 GCCATGTTA: TGAATCTGT AGCTTACAC TGAAGGAGC ATAGCATTA
701 AATGCTCTT: GCTTCTTTGA CAGTATGGG A

SEQ ID. No: 22

FIGURE 24

53/99

Session Name: crick.princeton.edu 1

Page 1

LL2-35(gel)

1. GACATTEUAG OCTACAGTAC ATATACAGTG ATCTGCTTC TOCAAGGATC
31. ATTGACAGG ATTCTCTGGT ACCTTCAGAA TGCTGCGCG CAGTGTATTA
101. ACAGCAGTGC CTGCTATGC TAGCGCTCA ACAGACAGT GAGTGGCGC
151. CTCAGAAAT CAATAAGAA GACTTGAGG GAAACAGAT GAGTGTGGT
201. CGAAGCTTC CCAAGATGG TTAGTACTGC TGCGCTGGA CAGCTTCAA
251. TTTCGCTTT CAGCTCTTG TACTTACG CAGTGGATC ATCATTTTC
301. AACGATAG NCTGACAG CAGTGTATT GATCTGTCA CTTACAGCT
351. CGAAGGACA TACATTTRA ATGCGCTTG CTCTTTGAC AGTAGTGGG

SEQ ID. No: 23

FIGURE 24

54/99

Unit 6 - tape

Session Name: crick.princeton.edu 1

Page 1

1 TGAAGTCCAC GCAATTAATG TGAATTAGCT CATTCTATAG GCACCTCCAGG
51 CTTTCACTTT TATCTTCCG GTCTGTATGT TGTGTGAAAT TGTGAGCGGA
101 TAACATTTT ACACAGGAAA CAGCTATGAC CATGATTAAG CCAAGCTCTA
151 ATAGTACTCA CTATAGCGAA AGCTGTAGG CCTGCAAGTA CCGTCCCGA
201 ATTCTCGGT CAGCCACGC GTCTGACGC TCGAGAGGA GCACAGAGG
251 GGGCAGACA GAGGACCTCT TCACTTAAG GAGACAGCA GCTTGGATA
301 GGAAGTCT' GCAGAGCTGG CAGAAAAGG AGATCAAGCA CTTTGAATAC
351 CTCATGTAC: TCACACACT GCGCGAAGG ACTTACAAAG ACTACATCA
401 GTATCTCTT TTTCTCTGG TCTCTCTCA CTACCTCTA GAGATCTTA
451 ACTTACGAA TCCCAAGCT TTTCTGATC TTTCTAAGC AATGCGGCT
501 CAGTCAAG: AAGGAAATG GAACTTACC CAGAGTTTA AGATCTTGA
551 AAGATGAA GCAGACATA CCGTCAAGT CCACTACTAC ACCCACTATT
601 CCTCAGCA: CATTCTGCT TCTACTTGG TCCGAATCC ACCATTCAAG
651 CAGGCTTCT: GCTCTTACA GGGCGAAGC TTTGATCTG CTGATAGAT
701 GTTCACTCT: TTAAGAGCA COTGCAATC TCGCTCAA GAGACATGA
751 GCGATGTA: GAGCTGACA CCGAATTCT TCTACTTCC GAGTTTFTA
801 ACCAAGTGA ATGCAATGA GTTCTCTTC ATGCAAGAT GAGCAGACT
851 GGGGAAATG CAGCTTCTC CTTGCTGTA GGGGATCC AGGAATTCA
901 TCACTTGA: CAGACAGCT CTGAAAGTG ACTTCTCA GAGCACTC
951 CAGCACTGA TAGACTAAT TTTTGGTAT AAGCAGCAG GCGCGCTTC
1001 TGTAGAGCA GTGAACACTT TCCAGCCTA CTTCTAGGT GATAGATAG
1051 ACCTGGCA: CATCACTGAC CCGCTGATCA AGAGCAGCAT CTTGGCTTC
1101 ATCAGCACT TTAGACAGT GCGCAAGCAG ATCTTCACTA AACCCACCC
1151 ATCAGTAA: ACCACAGGA AAGCCAGG GCGTGAAAG GATGCTTCA
1201 CCGCTGAG: CTTCTAGG CAGTCAAGT CTTCTCTCA CAGCTTCCA
1251 GCAGTGAAG CTTCTAGGT CAGCTCAA GATATGTACC TTTCTCTCT
1301 AAGCTGAA TCCCGAAG GGGCAGTG CAGATCTC CTTACTGAG
1351 AGTCAATCT GCGAGTGGAG AAGAACAGC TCTGATGCC CCGTCTCTG
1401 AACAGACT TCACTGAG CTTTGAATC TTAGTCTCT GCTGCGGAG
1451 CTAGGCTT GAGAGATCC TGAATACCT TGAAGCTG GCTGCGTGG
1501 GTCTCTCT GTGCTCTTA TCGCTTCC CAGGATAT CTTCACTCC

• DD16-1a pc

Session Name: crick.princeton.edu 1

Page 2

1551. GGGGCAACG CAATGGTGTG CATCTGGGAG CTGAGCCTGG TCAAAGGTGG
1601. CCGGAGAGGT GTGAAACTGC GACAGGCTTT GTATGGACAC ACTGAGGCGG
1651. TCACATGTCG GACAGGCTCT GTACGCTTCA GCTGCTGGT GAGGCGATCC
1701. CAGGATGCG. CTTGTATGCT GTGGGACCTG GACCACCTCT CTGCTGTGGC
1751. CTGCTGCTCT GTGACGCGGG AAGGCACTTC AGCCATTGCC ATCAGTGATG
1801. TCTGTGGAG: CATGTCTGCG TGTGCGGAG CCCACTTCTC CTTGTGGAT
1851. GTCAATGGAG: AAGGCTTGGC CAGTATTACC ACAGGCTGGG GCGGAGAGG
1901. AACCATAACX TCTGCTGCA TATGAGAGGG GCGAGGCTGG GATGCAAGCC
1951. ACGTATCTCT CACGCGGAGT AAGGACGCGA TGTTCGGAT TTGGAGGCA
2001. GAGGACTTCA AGATGCTCT TCCGAGGCG GCACTGATGG AAGAGGCTTC
2051. CAGGAGGCT: CTAGGCGGCA GAGGTCACAA GTGGGCGAG AATCTTGGCC
2101. TGAAGGAGG GCTGATGTC AGTGTGCTC TGAATGCGA GCGGAGGAG
2151. GCAATCTCTU CTGTGACAGC TCTGCGGATC ACTAGGAGCC AGAGCAAGCT
2201. CTTGTGTGG: GATGAGAGG GCGATCTTC TCTGCTCTG CTGATGGTA
2251. GAGGAGAGG: AGCTGAGGG ACCGAGCTGA AAGGCTGGAG CCGTGGGCTC
2301. GCGAGGAGG GCTGACAGG ACAAGATGAT GTGTAGGCTG GCTGCTTAA
2351. CCGAGGAGG TTTTGGGCG GCTGAGCTC ACAGGTTCT CAGGAGGCTC
2401. CTGATGCTT: GACGCTTGA CCGTAAACAT GTCTGTAGTC TATGGAGCTT
2451. CTGTAGAGG GATCTTGGTA GACACTGATG CTGGAACTG AGGCTGATG
2501. GGAATCTCT: AGTGGCTCA CTTACCAAG GCTTATTGCA CTGGGAAAG
2551. AAGAGGAGG: GGTATTGCT TCAATGAAAC CAGGCGACTG TCTTTATTT
2601. ATGAAACT: CATTTTGGG AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA

SEQ ID. No.: 24 (cont'd)

FIGURE 24

Session Name: crick.princeton.edu 2

Page 1

Hdd2.Seq Length: 499 August 12, 1997 14:21 Type: N Check: 4083 ..

```
1  CCCATTACTT CCACCTCCCA TGAGGACTGC ACCTCCGCCA GCGATTTTAT
51  GTGAGCTTG1 ATTACTGTAT GGAATGGAG GGTGAGAAA GTGTGGGAT
101  CATAAA-GAA GGCAGAACCA AAGACAGAG GGTCTTTGGG CCATGAGTCA
151  CAGCAAGGAA AGCTCTCCAC CAACATCTG CACTGGACTC CTACAGGAC
201  AAGAAATGAA CTATCTGTT CAGCCCCGA GACTGGGCT TACATGATT
251  TTTCATTTG CTACCTTACC TAATAGAGCA GAGATGTAAA TATATTTCTT
301  ATTTTAGAGT TGTGATGCT CAGCTGCAAT GGTGAGAAC TACTCTCAT
351  TTATTCTCTT CCAAGGCAAT AAGAGGAT GCACCAAGA CAGTCTGTCA
401  TCACATCTA TCAAAAGAGC TATGTGCA GTAGACTCT TCAAAAGAAA
451  AAGAAAAAC AAAAAAAG CATAAACAG ATGTCTTCTG CTTGAAAA
```

SEQ ID. No: 25

FIGURE 24

Hdd2.

Session Name: crick.princeton.edu 2

Page 1

14-A3-T7.Seq Length: 688 July 24, 1997 11:39 Type: N Check: 9884 ..

```

1  CCGGCTGCGC CCGGAATTAA GTGAGTCCAT TCTTGCACAC CGAGAAAATA
51  TCCGACGATG GATAGAGAGT TTTTCTGAAC CCTTTGGGAA GAGACTTCTT
101  CAGTATCTCT GATGCTAGAN GAGAGCTCA TAATCTAGA GCACATAATG
151  ATGGGAGAGA TCTTTGACT CATACAGATG TCACTCTTT CCAGACAATG
201  GACCAAAATG TGTCTAATAT GAGAACTAT ATCAGAAAT TACAAGAAA
251  CTTGCTGAA CTTTCAGTGG ATCCAAATGG ACATTCATTT TTTTCTCTT
301  CAGTTATAC TTTTCCAAA ATAGGAGATG AACCCGAAA GGTTTTCAT
351  GCTTCACTC AACTCTCTG AACTCTCTG AATTAAGG AAACCAAGAA
401  ACGCTGAGA GATTCTGACA GTGGACTAG AAAAATGCG TATTGCTCT
451  CATACTCATG AACCGAGCTC GTTCTCTTTA ATAAAGTCCN AAGAACCTAG
501  AAAGACTGGG AGATCAAAA AGTCCACCC AGGAANTTCA TCCCTTATGA
551  AATGAAGATT GATGCCCCCT GCTTTTCAA TGAAGGAAT GCCCAATTN
601  AAGGTTTTG AATTTTCAA ACCCCGAGC AACCAATCC TGGAAAACC
651  CTTGGAATG AAAAATTTT TTGGGCTTT AAAAN

```

1 (3;5) (325.1; 934) - vic gene, NPM-MLF!

SEQ ID. No: 26

FIGURE 24

Session Name: crick.princeton.edu 2

Page 1

34-D7.Seq Length: 651 July 24, 1997 11:37 Type: N Check: 6224 ..

```

1  CAAAGAGACG TGTGTTAGC AATATAAAC TGGCAGTGG GACAAGCTC
51  GCGCGGAAAT: GCTGTCACTA CTGCGGANTT CAGGAATGCC TGCTGTGTC
101  GAGGAGCAA GAAATGNNC GCACAGACAG NNGAAGGC CCGAGAGNC
151  GNTGCTCTN: GAAACCGAG AACACAGA NCCCTCTCC CCTTGGCCC
201  GTTTC/ND: GAACCTGAC ANCAAGAA NAGAGAGAN ACCCAGCAG
251  NCGGCAANT CTATGANTC CTGACTGCT CCAGGAAN CATCCCCGC
301  HGGCAGAAA AATCCCTGG GCGGCGAGA ACTGCCCCA GCGGACTAG
351  AACTGCTTCT TGAATGACN HTCTGAGAC TGTGTTCTN ACCAGTACA
401  NACAGAGAT GAGCATGAC ATGCTGCTT TCTCTGAT NCGGCGCTG
451  GGTATGNC ACAGAAAC AGCGGCGAG GAACCGAAA ANCGAAGAN
501  GCGGCGAGA TGTGTAATC HGGCGGAGC ACAGACTTT GAGGAGGAG
551  GCGGCGAGC GCGGCGCTN HTGTTGCTA CTGCGGCGG GAGGCGGCG
601  AATGCTGTC CCGGCGGCG CCGGCGGCG GCGGCGGCG TTTCTTTCT
651  N

```

TINUR = NGF1-B/cir 77 ceta type transcription factor

SEQ ID. No: 27

FIGURE 24

Session Name: crick.princeton.edu 2

Page 1

14-P4.Seq Length: 683 July 24, 1997 11:37 Type: N Check: 1900 ..

```
1  GCGTCTGTTT GGAATCAGA AACAAGGAGG TGATGTAGCC CATCTTTCC
51  TTTCAGATG CTAGAGGTGT TCTTCTCC ACCCAGACC AGGGGAGATC
101 CACTCTTCC TCCGCTTGA CCGTTTCT TCCGCTACC TCCCTCCCG
151 TCTGTTCC TTCTCTGAG ATGGTCATG CAGTTTAA GAGAACAGA
201 TCTTTATT TCGCTCCTG AAAAGCTAG GTACTCTCT TCACTGTTT
251 GAGCTCCAT TCTCTCTCT CATTACCTG CTCTTACTG ATAAATTTCC
301 TCGATTCTT AACTTTTCA TACATTTT TTTAACTACT ATATTGATT
351 TCGTTTAAA AGGAAAGTG CATTCTAT CAAATGTGT ATTCTCTTA
401 CGCTTTTCT TGTATACCA TTCTCTAGC TTATCTCTT TATATTTTA
451 GAGGAACTC CGATGTTAT GATCCCACT GTATGATTG TAAACAGAC
501 ATATGTAAG TCCCTTTTC AAAAAGAG GTGTGTTGA AATCATCCG
551 ATCTGCCAG GAACTGTCT CAAAGGAA CACTACTCT TCTGTTCTT
601 CCGTTTCTG TGATCATCG CAAAGAGTG CTCTCCCTG AATTTGCTT
651 TCGTTTGTG TACCAAAAA TTTTCTCTG TTT
```

JAP-kinase

SEQ ID. No: 28

FIGURE 24

Session Name: crick.princeton.edu 2

Page 1

38-A5-T7.Seq Length: 624 July 24, 1997 11:37 Type: N Check: 1627 ..

```
1  AATGACCTC CTTTATTC CACTTCGTA NICATTTGCG TTCTGATCTT
51  GTCAACCCG GCTGACCCG TTCTGACCG GCGATGGCTT CTTTACTAAG
101 CTTTCTTTT TAAGGAATG CTGTTTTTT TTGAGGTTT TCAAAACATT
151 TTGAAACCG TTACTTTTT TGACCACCA CCATGATTT TCAAAAAAT
201 CCGGGGTTG GTGGGTTTT GTTTTTGT TTAATTTTG GTTGGTTGC
251 CTTTTTTTT TAATGGGTT GCGCCCATG AATGGGTGC CCACTCACTT
301 CTCTGATG GAACGACTG TGAATGGCT CTTTGTGCA AACTGAACAA
351 MCTGTCCTT TTTTCCACT CCGTGTGAG TTTCGAATG TTATGTGCA
401 AGACCCCGG GGTGTGGCA CACTGCGCT GGAACCCCA CCCCTGCTC
451 CATCTGTC TGTGCGCGG CACATAAAA CATTCAACG TCCCTGAAAA
501 GTTCTTAAA AAATTAATT AATTTGTCG AATTTTACTG GAAAAAATT
551 AACTGACTC GATCCATTT TTTCTTTCN AATGGGCTGA AAGAAATCCT
601 TGTTCGCCA TTTTCTTTG UGCN
```

glycine-rich RNA binding protein

SEQ ID. No: 29

FIGURE 24

Session Name: crick.princeton.edu 2

Page 1

18-Q2.Seq Length: 641 July 24, 1997 11:37 Type: N Check: 656 ..

```
1  ATCCAGACC AGCCACAGG GTCCCTTGAG GTCTATAATG AAAGATCTGC
51  ATTCTGATGA CAATGAGGAG GAATCAGATG AATGGAAGA TAACGACAAT
101 GACTCTGAAA TGGAGAGGCC TGTGATAGA GCAACAGCC GAATGCCAG
151 AGTTAGCTTA AATGATGCA GCGATAGTGA AAGCAGTTCT GCTTCTTCAC
201 CCGTCATCA CGAACCTCCA CCACCTTAC TAAACCCAA CAACACCAG
251 ATTCTGAG TGAAAATCC NTAAAGCAA AGCAATCAG ATAACCAAT
301 AAGAAATGT CAATGTGACA AGCATACCT AGATGAAGTG GTNAGCTTC
351 CCGCAATTT NTCTACCTG AGAGAGAGAC ACATCTTCA GCGATCTTG
401 AACCTTATG AAGAACTGG GACACTTCA TATCACCAG ACAACATTG
451 ATTCTGATCT TTGCTGCTG GACAAACCA CAGTCCCTTA ACTACAGAT
501 TCTTGAAG: ATCTGAGCA TCTTGAAGAT ATAACCACTG GATGCTCTA
551 GAACCAATGT GTTTTSSSSS NNNNGGOTT NNNNNTTTG GGTGTGAATT
601 TTGCTCCX: TTGTTTAAAT GAAACCCCC CAAATGATC N
```

LTG9/MLLT3

SEQ ID. No: 30

FIGURE 24

Session Name: crick.princeton.edu 1

LL505

Page 1

1 GCGAGCCGAC GGAGNDOCTC TACGGTGGAT AGCGTUTTCG GGAAGCTATC
51 CCTCGAATTA GCGGAGTCAG GCAGAGAGGG GCGGCGGNGT GCTTCCGCCC
101 TTCTAGGAG GGGCTGCATT GCAUGGGAGA CCCAGCGGCA GATTCTGTCA
151 CAGAGGAGGG AGAAGGCGTG AGGAGACAAA GCGTCACAT GCGCGACAGC
201 TTCTTTCAGC AGCTCTCTCC TCTCCAGTCC AGAGCGGACC CCGGAGCCCC
251 TGAGGCATCC CTTCCTCTTT CGAAGACCC TAATCATCA TTCTAAGAG
301 GGGATCATGA GGGAGCCTGT AAGTAGCCAG TACAGCTCCT TTCTTTTCTG
351 GAGATGCCC ATCCACAAC TGGATCTTTC NGAGCTGGAA GCGCTGGGCC

SEQ ID NO: 31

FIGURE 24

63/99

Session Name: crick.princeton.edu 1

82-67

Page 1

1 ACCGCTGCGA GAGACGACA GAAGGGAAGC TAAGGCGTGG TGCGTGCCTC
51 TGAACGCCCG OTTGACGAA CGCTGTGCG CCCTTGCGG ACCTNAGCGG
101 AAAAGAGATG GCGGTGCGG TGCTGCAAGC TGTGCGCGG GTTCATCTTG
151 ATCTGACGC TTCTCTANTT TGTCTCAACC ATCTCTGAG CACAGAAAH
201 GAGAACTGA TGGGTCTGTG TATAGGCGAN TTGAATGATG ACATAAGGAG
251 TGACTCCNAA TTACATACA CTGGAACGGA AATGCGACA GTCCNAGAA
301 AGATGGATAC CATCAGAATT GTTCATATCC ATTCTGTCAT CATCTTGCGG
351 CTTCTGACA AGANAAANGA CCGTGTNAA ATTTCTCCAN ANCANCTGTC

SEQ ID. No: 32

FIGURE 24

Session Name: crick.princeton.edu 1

LL2-76

73

Page 1

1 GATGGCGGTC TCCACAGGAG TTAAAGTTC TCCTAATTTT CCGTTGTTCG
51 AAGAACTTGA AGAAGACAA AAGGAGTAG GTGATGTTAC TGTAGCTGG
101 CCGCTTGAAG ATGATGAAGA CATGACACTT ACAAGGTGGA CAGCATGAT
151 TATTGGGCCA CCAAGGACAA ACTATGAAA CAGAATATAT AGCGTGAAG
201 TAGAATGTGG ATCTAAATAC CCAGAGCTC CTCATCAAT TAGATTGTA
251 ACAAAAATTA ATATGAATGG GATCAATAAT TCCAATGGAA TCGTGATOC
301 ACGAGCATA CAGTATTAG CAAATGGA AAATTCCTAT AGCATTAAAG
351 TCATACTTCA AGANCTAAGA CTCTTATGAT GTCCAAGAA AATATGAANC

SEQ ID. No: 33

FIGURE 24

Session Name: crick.princeton.edu 1

1 ACCGCTGCGA GAAQACTACA GAAGCGGGAG CTCTGCGTC CCAAGGACTC
51 CAGTACACCA CCATGCGGGA TTCTGAGCOT CTCTGCGGCT CCGGCTGCTG
101 GTTAGCCTGC ACCAGCTTCT CCGGCACCAA AAAGGGAATT CTCTGTTTG
151 CTGAAATTAT ACTGTGCGTG GTGATCTTGA TTGCTTCAQ TGCATCTACA
201 ACATCGGCTT ACTCTCGCT GTGCGTGAAT GAGATGATCT GTGCTGCTGT
251 CTTACTTCTC TTCTACACOT GTGACCTOCA CTCCAAGATA TCATTGATCA
301 ACTGCGCTTG GACTGACTTC TTCAGATCCC TCATAACAAC CATGCTGTAC
351 CTGATCACCT CCATTGTTGT CCTTGTAGAA AGAAGAAGCA GCTCCAGAGT

76

B3-77

Page 1

SEQ ID. No: - 34

FIGURE 24

Session Name: crick.princeton.edu 1

C2-48

77

Page 1

1 TACCGCTCCG AGAAGACGAC AGAAGGGGCT CCOCCTTCTT CCCGACCTCC
51 GCGTCCCGCC GACACCGGTC TTTCAGTCC GCGACCCCGT GCGATCGTTA
101 CCCCCGTGCA CGATGCCCCC GAAAAAGGGA GGAGATGGAA TTAAACCCCC
151 TCCAATTATT GGAAGATTTC GAACCTCACT GAAATTCGT ATCCTTGGAT
201 TGCCAAATOT TGGAAATCT ACCTTCTTCA ATGTATTAAAC CAATACTCAG
251 GCTTCANCAAG AAAACTTCCC ATTCTCACT ATTGATCTTA ATGAGAGCAG
301 AGTCCCTGTG CCAGATGAGA GGTTCGACTT TCTTTGCCAG TACCATAAAC
351 CAAACAACAA GATCCTGCTT TCTAAATGT AATGGATATT GCTGGCCTTG

SEQ ID. No: 35

FIGURE 24

Session Name: crick.princeton.edu 1

(325

44

Page 1

1 ACCGCTGCGA GAAGACGACN GAAGGCTCC TTTTCTCTGC GCGCGTTTC
51 AGTCCCTCTT GAGTTGGCT GTGCTCNANT CCGTGGGANA CCGGCGACC
101 GCAGCCATGA CAGAACTGA TGTNAATCGG AAGGCTATC CCTTCNCANA
151 TNGCCACCTC ACCAAGAANC TGCTGGACCT TOTTNACAG TCTTGTNCT
201 ACAANCACT TCGAAAGGA CCAATGAACC NCCAAACCC TCAACAGAG
251 CATCTCTGAN TTCAATGTA TGCGCANAN CTGAACCTTG GAGATCTCT
301 GCACCTCCCT CTNCTOTCGA AAAAAAAT GTCCCTACCT NTTGTNCHN
351 CCAACAAGCT TTANGAAGG CCTNTGGTT TCCAGCCNTC ATCCCTGTCT

SEQ ID. No: 36

FIGURE 24

LL2-89

Session Name: crick.prima@etn.edu 1

Page 1

```
1  GAAAGGCACC TAGATACCCG GATACCTCCG GGTATCTCCG AGCTGGACTC
51  TCCTGTCTCTT OCCACATTTG ACTGAAGAAAG GACAACCGCG TCAAGGTTTG
101 AGGACTGACA GATTCTANAC CCAAGCTTCC TENDCCCCCA NAAAGCCACC
151 ANACCTCACA TGGAACTAAA GGCTCCATCC CCANCCCGCG TCCCCCAGCA
201 GGANANQAAA TTCCGTOTTC TTOTNRCOT CACTGGCAGC CTGCCCCCTC
251 TGAAGCTGCC TCTCTGNTA TCTANNCTOT TGGACGTTCC TNGATGTGG
301 AAGCGCCOTT CTGACCCGTT TCTCCACATT GACCTGCGGA TGTGGCTGA
351 CCTCATCTTA GTGGCTCCCC TCNATGCRNA CACTGTGNGG AAGT
```

SEQ ID. No: 37

FIGURE 24

69/99

Session Name: crick.princeton.edu 1

SBS-A56

51
Page 1

1 ACGGCTGCGA GAAGACGACN GAANCGGGGC AGGATGGCTC GGTGGTACAG
51 TTCAAGATCA ANAGCATAC CCCACTGAGC AAGCTGATGA AGCCTACTGT
101 GAGAGGCAAG GCTTGTCHAT NANGCAGATT CGATTCCGOT TTGATGACA
151 ACCAATCAAT GAAACAGACA CTCACCCCA NCTGAGATG GAGGATGAGG
201 ACACCTTTGA TOTATTCCAG CAGCAGACAG GAGGATCANC CTCGCCAGGG
251 AGGCTCCAC ACCCAACGOT TGTCTGACC TGTGCTATTG AGCAGTGACC
301 ATGCTNCCAG ACCCA

SEQ ID. No: 38

FIGURE 24

Session Name: crick.princeton.edu 1

C4-23

T0

Page 1

```
1  ACGGCTCCGA GAAGACGACA GAAGGGGAGA GCATCATGOC OCTAAGCGOT
51  CGACTGOCAT TGGGCGCGCT CAGACTGTGG GGTCCGGGAG CATGTATCTG
101 ANCTGCTTCG GACCAANAGC TUCAGTAAGG TGCTGTTCAT GGTGGGGGCG
151 GGCATCANCA CAGCCAGTGG CATCCGGGAC TTCAGATCCC CAGGGGAGCG
201 GCGTCTACAG CAACCTTCAG CAGTATGACA TCCCGTACCC TGAAGCCATC
251 TTGGAAGTTG GCTTTTTCIT TCACAACCGC AAGCCCTTTT TCATGTTGOC
301 CAAAGACTOT MCCCTGNGCA CTACAGGCCA ATGTCAGTCA CTACTTCTG
351 AAGCTGCTCC AGGACANOGA CCGCTTCTGC GCTCTNTACA CANAAATCGA
```

SEQ ID. No: 39

FIGURE 24

Session Name: crick.princeton.edu 1

LL5-68

52

Page 1

1 GGCATCAAN COTGCTTTG GTCATGGCG ACTGCTCAA GATGGACCC
91 TCTACCCCT CAACACTAG GCAOCCAGT CTGAGGGAA AGAGGCTGT
101 GAGTCCTCT GGTGTTTACC TCCGCTGCA CTGCTCCAC ACCGAAAGA
131 CAAGCCGTT GCTGAACCA TCCCATCTG TAGCTTCTGT CTTGGTACAA
201 AAGAACAGAA CCGGAGAA GAGCCGAGG AACTGCTCT CTGCGGAGC
251 TCCGCAATA NCGGTCATC ATGCTGTTA AGTTCTCCC CANAGCTNAC
301 AGTGAGAGT AAGCCTTACN GTGCTTACA TTGANTGAA AACATCACT

SEQ ID. No: 40

FIGURE 24

>SC|LLS-96-full.seq|14350
CCCACGCGTCCGATCTCCTCCAGGGCCACCAAGCACCTCTGAAGAGCCATGTTCCAAGCTGCCGGAGCCGCCAGGCCAC
CCCCTCTCATGAAGCCAAAGGCAGCAGTGGCAGCAGCACGGTACAGCGGTCTAAGTCCTTTAGCTTGCGGGCTCAGGTGA
AGGAGACCTGTGCAGCCTGCCAGAAGACTGTGTACCCGATGGAGCGGCTGGTGGCAGACAAGCTCATTTTCCACAACCTCT
TGTTTCTGTTGCAAACTGCCACACCAAACTCAGCCTGGGCAGTTATGCTGCAATGCACGGTGAATTTTACTGCAGACC
TCACTTTCAGCAGCTGTTTAAGAGTAAAGGCAACTACGATGAAGGGTTTGGTCGTAAACAGCACAAAGGAGCTCTGGGCCC
ACAAGGAGGTGGACTCAGGCACCAAGACGGCCTGAGACCCCTTTAACACCCATTCCCTCCCAGCACATGGCCTCCCGCTG
GGCAGTGGAAAGGAGATTAACCCGGGGGCGGGGGTGGGAGAGGATGAGGCTCCCTCACACAGGTTTCAGGCATAAGGCT
CTGCTCCAGGATTCCTTACTTTTCCATGGGAGGTTGGCGTTGGGAACCAAGAAATTTTACCATACTGTGTCCTT
TAGTCCACCTCATCTACCCACGGCTCCCTGGGAGGCCACAAGCCCAGCTTCCATACTTAGGTGCTTTTCTCCAGCAA
GGAGTCAGCATGCCTCCTCAGGGTCCCAAGCTCCCTCACTGCCACCTGGGCCTTGTGTACCCCTTGTCTCCCATCTA
CCTCTGCCCCCTTAGCCTGGTAATGAGCCACAGAGACTGGAAGAGGGAGAGTGCCATNTACTGGGCCTCATAGATGCCACC
TCGCTGAGGGGGGAGGGCTGGGAAGAGGCAAGACAGCCTGCAGCCTTCAGGGTCTGGGGGTCCCTTGCACCACAAAGCT
AAAGCTCTTGCTAGAGCCTCAGCTGACAGGGTCGGCAGTAGCTATGCTCCTCCATCTGTTGTGCTGTTCTGTTGTGATCA
ACCCTCTTTTAAAAACATTTAAACAGCAAAAAAAAAAAAAA

SEQ ID. No: 41

FIGURE 24

>SC|LL5-96-full.seq|14350 Frame 1
MFQAAGAAQATPSHEAKGSSGSSTVQRSKSFSLRAQVKETCAAC
QKTVYPMERLVADKLI FHNSCFCKHCHTKLSLGSYAAMHGEFYCRPHFQQLFKSKGNYD
EGFGRKQHKELWAHKEVDSGTKTA*DPFNTSLPAHGLPLGSGKEINPGARGRG*

SEQ ID. No: 42

FIGURE 24

>SCISA7-full
AGCTCTAGTC CCCAGAGATG TCGCCACTAC TGCTGCTGCT GCTGTGCCTG
CTGCTGGGGA ATTTGGAGCC TGAGGAGGCC AAAGTATCC GTGTCCCTCT
TCAACGAATC CACCTTGGAC ACAGAACTTT AAACCCACTG AATGGATGGG
AACAGCTGGC AGAGCTTTCT AGGACCTCCA CCTCTGGTGG CAACCCCTCC
TTTGTGCCTC TCTCCAAGTT CATGAACACC CAGTATTTTG GAACTATTGG
TTTGGGAACG CCTCCTCAGA ATTTACCCGT TGTCTTTGAC ACGGGTTCTT
CCAACTTTGTG GGTTCCTGCC ACGAGATGTC ATTTCTTCAG TTTGGCATGC
TGGTTTCACC ATCGCTTTAA TCCCAAGGCC TCCAGCTCCT TCAGGCCCAA
TGGGACCAAG TTTGCCATTC AGTATGGGAC CGGGCGGCTG AGCGGAATCC
TGAGCCAGGA CAATCTGACT ATCGGGGGGA TCCACGATGC TTTTGTGACA
TTTGGAGAGG CTCTGTGGGA GCCCAGCCTG ATCTTTGCTT TAGCCCACTT
TGATGGGATC CTGGGCTCG GCTTCCCCAC TCTGGCTGTG GGCGGAGTTC
AGCCTCCGCT GATGCGATG GTGGAGCAAG GGCTGCTGGA GAAACCCGTC
TTCTCCTTTT ACCTCAACAG GGATTCTGAA GGGTCTGATG GGGGAGAGCT
GGTCTAGGG GGCTCAGACC CCGCTCACTA CGTACCTCCC CTCACCTTCA
TACCAGTCAC CATCCCTGCC TACTGGCAGG TCCACATGGA GAGTGTGAAG
GTCGGCACAG GGCTGAGCCT CTGTGCCCAG GGCTGCAGTG CCATCCTAGA
CACAGGCACA TCCCTCATCA CAGGACCTAG TGAGGAGATC CGGGCCTTGA
ATAAAGCCAT TGGGGGATAT CCCTTCCTGA ATGGGCAGTA CTTCAATTCAG
TGTTCCAAGA CGCCAACGCT TCCCCTGTC TCCTTCCACC TTGGTGGAGT
CTGGTTTAAC CTCACAGGCC AGGACTATGT CATCAAGATT CTTCAGAGCG
ATGTTGGCCT CTGCTGTTG GGCTTCCAAG CTTGGATAT CCCCAGCCT
GCGGGACCCC TCTGGATCCT TGGGGACGTC TTTTGGGGC CCTATGTGGC
TGTCTTTGAC CGTGGGGACA AAAACGTCGG CCGCGCGCTG GGAAGTGGCG
GTGCTCAGTC TCGTTCAACA GACCGGGCAG AAAGAAGGAC TACGCAGGCG
CAGTTCTTCA AAAGACGCCC TGGTTAGGGT ACAAGCTCAC CGGGCCACAG
CAGCTATGCT TCTTTCCAAT TAAACAACT AAAAAAAAAA AAAAAAAAAA
AAAAAAAAA A

SEQ ID. No: 43

FIGURE 24

>SCISA7-full Frame 3
MSPLLLLLLC LLLGNLEPEE AKLIRVPLQR IHLGHRILNP LNGWEQLAEL
SRTSTSGGNP SFVPLSKFMN TQYFGTIGLG TPPQNFTVVF DTGSSNLWVP
STRCHFFSLA CWFHHRFNPK ASSSFRPNGT KEAIQYGTGR LSGILSQDNL
TIGGIHDAFV TFGEALWEPs LIFALAHFDG ILGLGFPTLA VGGVQPPLDA
MVEQGELLEKP VFSFYLNKDS EGSDGGELVL GGSDPAHYVP PLTFIPVTIP
AYWQVHMEsV KVGTLGLSLCA QGCSAILDTG TSLITGPSEE IRALNKAIGG
YPFLNGQYFI QCSKTPTLPP VSFHLGGVWF NLTGQDYVIK ILQSDVGLCL
LGFOALDIPK PAGPLWILGD VFLGPYVAVF DRGDKNVGPR VGLARAQSRS
TDRAERTTQ AQFFKRRPG

SEQ ID. No: 44

FIGURE 24

76/99

>SC|DD136|14044
CGGTGACGCAGGACCAGGACTCGCGCGTCCAGCGGAGAAGCAGGAGAAGCCGGCGACCTTGGCTCTCAGCCTGATCCCT
GTCTTGGCGGCGCTGAACATTTCGACGCTGGAGAGATGGCGTTTCGTGAAGAGTGGATGGTTACTTCGGCAGAGCACCATTCT
GAAACGCTGGAAAGAAGATTGGTTTCGACCTGTGGTCAGACGGTCACCTGATCTACTACGATGATCAGACTCGGCAGAGCA
TAGAGGATAAGGTCCACATGCCCGTGGACTGCATCAATATCCGCACGGGGCATGAGTGCCGGGACATCCAGCCTCCAGAT
GGGAAGCCCGAGAGACTGTCTGCTGCAGATCGTTTGGCGAGACGGGAAGACCATCAGTCTCTGTGCAGAGAGCACAGACGA
TTGCCTGGCATGGAAGTTTACACTGCAGGATTCCAGAACAACACAGCTTACGTTGGTTCAGCAATCCTGTCTGAAGAGA
CTGCAGTGGCGCGCTCCCGCGCTCCCTACGCAACCTATGCTACACCNACCCCTGAGGTCTACGGCTATGGTCCATACAGC
GGCGCATACCCCCGAGGAACCTCAAGTTGTCTATGCCGCCAACGGGCGAGGCATATGCAGTGCCATACCAGTACCCGTATGC
AGGAGTTTATGGACAACAGCCTGCCAACCAAGTCATCATCCGCGAGCGGTACCGAGACAATGACAGTGACCTGGCTCTGG
GCATGCTCGCCGGGGCAGCCACCGGCATGGCCCTGGGCTCTCTGTTCTGGGTCTTCTAGAGCCTTCAACATTTCTGTGC
ATAGCTTCTGTAGTCTCTGTGTCAGTAATTTGATTTCGAGGGCATTCTGTTTGTGACAAGTGTCTTTCATAATAATTT
AAATAGTTCTTTTGAAGGTGGTAATCTAATAATTTGTGACTGACCTGCATGGTACCACAAAGAAAGCCCGAGGTATGCTGT
GAGTGAGAGCCTGAGTCCTTCGGGTACTAGCTTGACCAAGTCTTTCTAGGGACTTTTGGATGGCTTTATGTAACAC
ACCCAGTTAAATGGGCAATTTCCGTCCAGTTAGGTGCAGTGTGAATTAAGGGATGGCTTTCCTTGTATGCCAATACTA
ATAGTCTGATGGAGGAAGATGTGTGCAAGTGTGGTGAGGAGAGTCACAGCTTCTTAACTGTGGATTCTCTTCTAGACC
CCTGCTGCGGTGTTACCTAGGAGCTGTGGGCTGGTGGCTCCTGCAAGACTATGGTGTGAGGACCTGTAACGTACCTCTT
GGAGCACTTAGGTACCCCTTGAAGCTCCTAGGTATCACCAGCAGGATTGGCTGCTCAGGATGCAGAGGGCCACCCCTCC
CTTTAAAAATTACGCTCCAGTAATCTGCCAGTTTTATTTCTTGTATTCTTCTGTTGCTTTTCTGGGGATGATTGG
CATTAGTCTGGAGTTAGGAATGATTTCGAGTGCCGGTGGGTGGAGGCATGCAGGGAGCTGTCCAGCGACCTGCTCTCAGT
GTTTGTGTTTAGGTATATTGATTGCCAGCTCAGGCTGCAGAGAGCCTATAGAGACTATTTTCTACTTGTAAGAAAGTAT
GGTGAGGGGAATTGAGGAGAGCCTTGTGTAATGTTCTGCTCAGGCCTCCTGGGGCCACACTGCGTGGTCTGGGGA
GGCTTTTCTCAGCAGTGGAAGGGAGGCCCGTGGCTGTCCAGAGTCTCAGGTTTGTAGTGAGAAATGGGATTGGGTAGA
GCATCTCAGGGATTTGTTCTAATCCCTCATGTTATGGGGATCCAGCCGTGTTCTCAGTCCAGACCCGCTCACCTCAGAAG
AGCTTAAACATTTCTGGTCCCCAAATGTGTGGCACTCTGAGAAGCTCACAATCTGGCTTTCTAACGAAAATTTGTATTT
CTAAATTAGAGAATACATGTTCCACGCATTTAAATTTATGTTCTTTTCTGTTTAAAGCTCCCAAATCCAGCTTTGTG
ACTGGCATATTTTAGTTTCAACAGTACCCCGGCACAAAGTGGGATGGCAGTGAAGGCCCCCGCCCTCTACTTTGC
ATAGTCTTGTGTTCTCCAGGGTGTCCAGGAAGCATTCACTCTGACTTTGCTCAGCCAGTGCATGCGTGCTGCTTGGC
GCCGTGCTGCTGGGTAGCTCTTCTTGGTCAGATCAAGTCTTCAACAGATCTCCATGTGAGACAGTTGCCAAGTAGATGA
GGTGGTGCCCATAGTGCTTTCTCGATCTCCTTGGGGACCTGTTGACACCTGCCATTTCCAGCTGACATTTGTTTTCT
GTCATCTCTGATAGATGGATATGTGACAACATGGTACGGACGCCGTTCAAGTGTGCTTTAATAAGCATGATGCTGATTT
TACATCCTGTGCTGTATGACTGCCATTTGCTCACAGTGTACCATTGCTAAAGCTCCGTGCTTACTTACAAAACACTAA
AACCAGTGGTTAGTGTTCACAGTGATTTTAAATTTTAGAGTTAGTTACTGGCATTCTTAAAGCCATAGAGTACTGAGTCA
CATCCCTGAAGTACTTTTGAACAGAATTGTCTCCTACTGTCCCATGGGTGTGCCCTGCCTGTCTCCTGGCCCCAATGGG
GCTAGCTGTACCAGGCAGCCATAGTTGAGCCTGATCATTCTGTACCAGTTTGACTTGATTATATACCAGAATGGAAT
ACATTCTTGGGCATCTCAGTTCCTCAGCCCTGATCCTCATAGACGCCACCCTTTCGATGGCTTTTGGCGGCTCACTTGTA
CCTCAGTGAGTCTCGGATTCTTGAGTTAGAGGGGACGACTTGTCCAGCATTGAGGAACATGTCTCCTCCACTGAGACTT
AAATGATGATGCAGGGCTGGAAGAGSCTGGCTGCTGACACTGCATCGTGGCTGATGTCATTGCTCTCCTAGTTCTTTGAT
TTAAGAACCCTTTCATATGGAAGGCCCTGAGGCTCCCTCAGATCGTCCCTTGCCAAGAAGGCCCTGGCTTAGGTCTAGTGC
CCACAGTAGCCTTCTGGAGTGTAGCTAGTTCTGCTGCTTTGAGACAGAATGGTTCAAGTTATTTCTACATCTGTTGTTG
ACCCCATGCACCCTCTCATTTTGCCTTCCAGTCTACGTAGATGAAAGATGAAAGGCAGAGGATGCAGACAGTCTTCTTTG
TGATTGCTTCTGTTATTTCTGTTGCTATACCAGGCCGCTTTCTCCTGTCTGTGSCATACAGTATGTTTATAAGTGAAC
TGTTAAATATTAATGATCACTAACCTAAAAA

SEQ ID. No: 45

FIGURE 24

77/99

>SC|DD136|14044 Frame 3
MAFVKSGWLLRQSTILKRWKXW
FDLWSDGHLIYYDDQTRQSIEDKVHMPVDCINIRTGHECRDIQPPDGKPRDCLLQIVCRD
GKTISLCAESTDDCLAWKFTLQDSRTNTAYVGSAILSEETAVAASPPPYATYATXTPEVY
GYGPYSGAYPAGTQVVYAANGQAYAVPYQYPYAGVYQQPANQVIIRERYRDNDSDLALG
MLAGAATGMALGSLFWVF*

SEQ ID. No: 46

FIGURE 24

78/99

>SC|DD116|14045
CACCACAGAGGACCCCTTCAACCTAAGGAAACACCCAGGCTTCGATAGGACCATGCTGCAGAGGTGGCAGAAAAGGGAGA
TCAGCAACTTTGAATACCTCATGTACCTCAACACACTGGCCGGAAGGACCTACAATGACTACATGCAGTATCCCGTGT
CCCTGGGTCTCGCTGACTACACCTCAGAGATGTTGAACTTGACGAATCCCAAGACTTCCGGGATCTTTCTAAGCCAAT
GGGGGCTCAGACCAAGGAAAGGAAGTTGAAGTTTACCCAGAGGTTTAAAGATGTTGAAAAGATTGAAGGAGACATGACCG
TGCAGTGCCACTACTACACCCACTATTCTCAGCCATCATTTGTCGCTTCTACTTGGTCCGAATGCCACCATTACGCGAG
GCCTTCTGCTCCTTACAGGGCGGAAGCTTTGATGTGGCTGATAGAATGTTCCACAGTGTAAGAGACAGCTGGGAGTCTGC
CTCAAAGAGAACATGAGCGATGTGAGGGAGCTGACACCTGAATTCTTCTACCTGCCCGAGTTTAAACCACTGTAATG
CAGTGGAGTTTGGCTGCATGCAGGATGGAACGACACTGGGGGATGTGCAGCTTCTCCTCGGGCTGACGGGGATCCGAGG
AAATTCATCAGCTTGCAAGACAGGCTCTGGAAAGTGACTTCGTGAGCAGCAACCTCCACCCTGGATAGACCTAATTTT
TGGGTATAAGCAGCAGGGGCGGGCTGCTGTAGAGGCAGTGAACACTTCCACCCCTACTTCTACGGTGATAGAATAGACC
TGGGCAGCATCACTGACCCGCTGATCAAGAGCACCATCCTGGGCTTCATCAGCAACTTGGACAGGTGCCCAAGCAGATC
TTCATAAACCCACCCATCCAGAAACACACAGGGAACCCAGGGCTGGAAAGGATGCTTCCACCCCTGTAGGCTT
CCCAGGCCACTCACAGTCTTCTCCTCCACAGCTGCCAGCACTGAGACCTCTCAGGTACAGTCAAAGATATGTACCTTT
TCTCTTAGGGTCGGAATCCCCAAAGGGGCCATCGGCCACATCGTCCCTACTGAGAAGTCAATCCTGGCAGTGGAGAAG
AACAAGCTGCTGATGCCCCCTCTCTGGAACAGGACCTTCAGCTGGGGCTTTGATGACTTCAGTTGCTGCCTGGGAGCTA
CGGCTCTGACAAGATCCTGATGACCTTTGAGAACCTGGCTGCCTGGGGTCCCTGTCTGTGCGCTGTATGCCCTTCCCCA
CGATGATCGTCACATCCGGGGCCAGCGCAGTGGTGTGATCTGGGAGCTGAGCCTGGTCAAAGGTGCCCCGAGAGGTCTG
AACTCCGACAGGCCCTTGATGGACACACTCAGGCGGTACATGTCTGACAGCCTCTGTACCTTCAGCCTCCTGGTGAG
CGGATCCAGGATCGCACTTGATCTCTGTGGGACCTGGACCACCTCTCTCGTGTGGCCTGCCTGTCCACCGGGAAG
GCATCTCAGCCATTGCCATCAGTGATGTCTCGGGAACCTTGTCTCTGTGCGGAGCCACTTGTCCCTGTGGAATGTC
AATGGACAGCCCTGGCCAGTATTACACAGCCTGGGGCCAGAAAGGAACCAACGTGCTGCTGCATAGTAGAGGGGCC
AGCGTGGGATGCAAGCCACGTGATCATCAGGGGAGTAAGGACGGAATGGTTCGGATTGGAAGACAGAGGACGTGAAGA
TGCTGTTCAGGAGGCGAGTGTGAGGAGGCCCTCCACGGAGCCCCAAGCCCCAGAGGTCAAAAGTGGGCCAAGAAAT
CTTGCCCTGAGCCGAGAGCTGGATGTGAGTGTGCTCTGAGTGGCAAGCCAGCAAGGCAAGTCTGCTGTGACAGCTCT
GGCCATCACTAGGAACACAGCAAGCTCTTGTTGGCGATGAGAAGGGCGAATCTTCTGCTGGTCTGCTGATGGGTAGGA
GACAAGGAGCTGGAGGACCGACCTGAAAGCCTGGAGCCCTGGGGTCCGCAGCAACAGGCTACAGGCACAAGATGATGTG
TAGCCTGGGCTGCTTAACAGAGCAAGTTTTGGGGGGCTCCACTCCACACAGTTCTCAAGGAGTCCCTGATGGTTTGCA
CCGTGTACCCCTAAACATGTCTGTAGTCTATGGGACTTCTGTAAGAAGGATCTTGGTAGACACTGATGCTGGAACTGACG
CTGCATGGGAAATTTCTACGTGGCTCACTTCACCAAGGCTTATTGCACTGGGAAAAGAGAGGGGTGGGTATTGGTTCA
TGGAAACACCCCACTGTCTTATTTATTAATACTCCATTTCCGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

SEQ ID. No: 47

FIGURE 24

79/99

>SCID0116114045 Frame 2
TTEDPFNLAKHPGFDRTMLQRWQKREISNFEYLMYLNNTLAGRTYNDYMQYPVFPWVLADY
TSEMLNLTNPKTFRDLSKPMGAQTKERKLKFTQRFKDVEKIEGDMTVQCHYYTHYSSAI
VASYLVRMPFFTQAFCSLQGGSFVADRMFHSVKSTWESASKENMSDVRELTPEFFYLPE
FLTNCNAVEFGCMQDGTTLGQVQLPFWADGDPKRFISLHRQALESDFVSSNLHHWIDLIF
GYKQQGPAAVEAVNTFHPYFYGDRIDLGSITDPLIKSTILGFISNFGQVPKQIFTKPHPS
RNTTGKNPGPGKDASTPVGLPGHSQSFLHSLPALRPSQVTVKDMYLFSLGSESPKGAIGH
IVPTEKSI LAVEKNKLMPPLWNRTFSWGFDDFSCCLGSYGSDKILMTFENLAAWGPCLC
AVCPSPMTMIVTSGASAVVCIWELSLVKGRPRGLKLRQALYGHQAHTCLTASVTFSLLVS
GSQDRTCILWDLHLSRVACLPHVREGISAIASDVSGTIVSCAGAHLSLWNVNGQPLAS
ITTAWGPEGTITCCCIIVEGPAWDASHVITGSKDGMVRIWKTEDVKMPVPRQAVMEEPST
EPLSPRGHAKNLALSRELDVSVALSGKPSKASPAVTALAITRNQSKLLVGDEKGESSA
GLLMGRRQGAGGTDLKAWS PGVGSNRLQAQDDV*

SEQ ID. No: 48

FIGURE 24

>SCIDD64|14046
CAGAGGCAGGCAGGACAGGCACTGCASGCACTGAGAGCCTTGCACTCAGAAGGGGAAGCACASACAGGAAGAGGCAGTGAG
GAGAAAAGGAGCASAACCTGAGACAAGCAATGGAGATCCCTGAGGAGTGCTGGCCACGGCTGAAAGTGTCCCTAAAAGACA
TCACTGAAATAATGGAGAGACATCTCAGTCACATGGAACGGACCCCTGTCTCACAGTCAAAAGCTCTCAGATGGAGACCTG
GTAAGATGGGCATCTGGAGGGCTGGTCTGCAGGGAATTTATAAGACCAACCACCCAAGAAGCCTGATTGAGAAGAGAGA
AGAGCTACTCAGTGTCCCTAAGCAGTTCTCACTTGCTGGCCAGTACATGGCAGAGATAAAAACAAAGGAATTTTCAT
CTTTTCAAGAACAAGCCATGTTTACACAGACTATAGAGAGGGTGGGTTTCAGTTCAACCCCTTAGTTAAAGGAGAAGGT
TGGGCACTTAGTCTAGAAGCTGGTATGGGTGACAACAACAGACAGAATCTGAAGATAACTACCACTCCCATTCAAAGCA
AACTTATTTTGTCTCAGCCAGGTTGAGTACATCCCATTTGGCCACCTGCCATTTTACATCAATGATCTTGAACCTCTCCC
AGGCTGCTCTCCAGGAACATAAAGTATTGAAGAAATCCTGGAGCAGACTACACACCACCGAGATGGACTACCCCTTACTG
AGGCACAGGGCTAAAAACTTTTCCACAGGTTTGGCTCTCATGCTAACCAAGGCCCTGTGCACCTGGGGGAATCTACTG
CTGGAAGCCATTTGAGAAGGTTTCAAAAGTGAGCACTTGGCTGATGTAAAGCAGCAAGCAGAAGAGTCTTTGAATATTT
ACATTATGGGCAGTTATAGTGGCTTTGGAGTTAAAGTTGGTGCGAGTGTAATATAACAAATTCAAAATCAAAAACAGCA
TTTTACAGTAAAACTCATCTAACTCGCAAACCAAGGTACAACATATCTGTAGCCAAAGATAGGTGGACCAGCAGAAGCAGA
TGGAAATGGCCAGTGGACAGCTGGCCTTGTAGCTAGCAATCAAACCTGGTCTGTTATTGATAGGAACTCGAGTTGGTAC
CTATTTGGGACATTATCCTGTCCAGTCAACAGAACTGAATTTAAGAAATGCTCTTCAACTGGCTAACTGCCTCAAAGACCAC
TACACTGCTCTGACTGAACTAGCTGCCAGATTCAAGAAGGGGAAGAATTTCTGACTGCTAGAAAAGAAGCTAAGCTTTT
CCTAAAGAATGTGAAAGGCTGGGAGGTTTCTGATCCTGAAGAACAGCTTAGGAAGTTAGTAGATTTTATGCAACATTGA
GTCAAAAAATAAAAGTTATGACATTTGGATTAAACACATGCCTCATAGATTGGGATCTGCAAAATTTTCTAATAAACATT
GTCAACTTCTGCAAAAATTCACCCACTTATAAACTCACTTTATTAAATCTCAGTTGTGCAGCCTTCTAGAACCTCATGT
CTACAAAGTGACAACTTTCTCAGGCACAATGCATCATACAGTGGATCAATCAGTCAGAGTCAGAGGAAGAACTAGTCA
AAATTACCTCATTTTCTGAATTCATTAAACACCTTAAAGAAAACCCACAAATACCTAATGGAAGAGAGTTTCAAACTGAG
CCCCAGAAAGAGTGGAAGAGCAAGAGAAATGGCTACATATGAAGTCACCACAGCTCTCAGCTCCTTCTTGAAGTACCT
CAGAGAAACACAGCAGCCAGACATGCAGCTGTTGCTACTCTCCATTGCTACTGGTGTAGGCTATCAGTTGGTAAACAGTA
TTTTTCAGCATCTTCTGGGGTGTGATGAGTTAACTTCTCTTGGATCAATGGAAAATAACGAACATAAATACCAAGAA
CTGAAAATATTTGCAATTACAGAGCCCGAGGCTTCTTGGTGCTCACAGCCCTAAGAGCCACAGTTGAAATCACAGATGT
TTCTACAGAAGAGAAAGGACAACGTTTGACATTAAATACAACAACATATGGGGTCACTGTTGTCTGAAGAAGTTGCACATG
TCCTCACAAAACATGGAGAACATCATGACTGGGAAAGGCTGGAGAATGATTTGAGATTACTCATTGAGGGGGACTATAAA
GCCACCACCCATTCTTACAAATGGATGAAGTAAAAAACAATTGCAAGTTTNTGCCATGAAAAGAAACAGACTTATAA
ACAACAAGGTAATGAAAACAGAACAAAAGAAATGATAGAAAATGGACATTTCTGGACTTACTCCAACGTTTAGGCCCTAG
ACAATTACTATCCAAAAA

SEQ ID. No: 49

FIGURE 24

81/99

>SC|DD64|14046 Frame 1
QRQAGQALXALRALQSECKHXQEEAVRRKGAXLRQAMEIPEECWPTAEVSLKDITEIMER
HLSHMERTLSHSQKLSGDGLVRWASGGLVLQGIYKTNHPRS LIQKREELLSVPKQFSLAG
PVHGTEIKTKEFSSFOEQAMFTQTIERVGFSSSTPLVKGEWGLSLEAGMGHNKQTESEDN
YQSHSKQTYFCSARFSYIPLATCHFHINDLELSQAALQELKSIEEILEQTTHHRDGLPLL
RHRAKNFFHFRFGSHANQGPVHLGGIYCWKAISEGFKSEHLADVQQAEE SLNIYIMG SYS
GFGVKVGASVNITNSKSKTAFYSKTHLNSQTKVQLSVAKIGGPAEADGIAQWTAGLVASN
QTWSVIDRKLQLVPIWDIILSSHRTEFKNALQLANCLKDHYTALT ELAAQIQEGEEFLTA
RKEAKLFLKNVKGWEVSDPEEQLRKLVD FMQTL SQIKSYDIWINTCLIDWDLQNFLINI
VNFCKNSTPYKTHFIKSQ LCSLLEPHVYKVTNFPQAQSIIQWINQSEEEELVKITSFSE
FINTLKKTHKYLMEESFKTEPPERVEEAKRMATYEVTTALSSFLKYLRETQQPDMLLLL
SIATGVGYQLVNSIFQHLLGCDELN FLDDQ MENNEHKYQELKNICNYRAQAFVLVTALRA
TVEITDVSTEEKQRLTLIQQHMGSL LSEEVAHVLT KHGEHHDWERLENDLRLLIEGDYK
ATTHSLQMDEVKKQLQ SXCHEKKQTYKQQGNENRTKEMIENGHFLDLLQRLGLDNYYP

SEQ ID. No: 50

FIGURE 24

82/99

>SC119114042
TCTTCGAAAGCCGGGCTGAGGGGAATCCTGGACAGGGGAATCCTGGACGTGGAGATCGTGAGTCATGGCTGCTTCCCGAG
ACGCTGATGAGATCCACAAGGACGTTCAAGAACTACTATGGGAATGTACTGAAGACATCTGCAGACCTCCAGACTAATGCT
TGTGTACGCGAGCCCAAGCCGGTCCCCAGCTACATCCGGGAAAGTCTGCAGAATGTACACGAAGACGTTAGTTCGAGGTA
TTATGGCTGTGGTCTGACTGTTCCCTGAGCGGCTGGAAAACGCCGAATTTTGGATCTGGGTAGTGGGAGTGGCAGGGATT
GCTATGTGCTTAGCCAGCTGGTTGGTGAGAAGGGACATGTCACCGGAATAGACATGACTGAGGTCCAGGTCCGAAGTGGCT
AAAACCTATCTTGAACACCACATGGAAAAATTTGGTTTCCAGGCACCCAATGTGACTTTTCTCCACGGCCGCATCGAGAA
GTTGGCAGAGGCTGGGATCCAGAGTGAGAGCTATGGTATTGTATATCCAACCTGTGTTATCAACCTTGTTCCTGATAAAC
AACAAGTCCCTCCAGGAGGTCTATCGAGTGCTGAAGCACGGCGGGGAGCTCTATTTCACTGACGTCTATGCCAGCCTTGAA
GTGCCAGAAGACATCAAGTCGCACAAAAGTTTATGGGGGAATGCCTGGGAGGCGCTCTGTACTGGAAGGATCTTGCCAT
CATTGCCCCAAAAGATTGGGTCTGCCCTCCACGTTTGGTCACTGCCGATATCATTACTGTTGAAAAACAAGGAGCTCGAAG
GGGTCTTGGTGACTGTCGCTTTGTGCTGCCACATTTCCCTCTTCAAACCTCCCTAAGACAGAGCCAGCCGAAAGATGC
CGAGTTGTTTACAATGGAGGAATCAAGGGACATGAAAAGGAACTAATTTTCGATGCAAATTTACATTCAAGGAAGGCCA
AGCTGTTGCAGTGGATGAGGAGACGGCAGCTGTCTGAAGAACTCACGTTTGTCTCGGATTTTCTCTTACACCTGTTG
ACGCCCTCGCTGCCAGCTCCCCAGGGGCCGTTCTGAGTTAGAGACAAAGGTTCTAATCAGAGATCCATTCAAGCTTGCAGA
GGACTCTGACAAGATGAAGCCCCAGACATGCACCTGAAGGCACGGGAGGCTGCTGTGGCAAGAGGAAAACTGCTAGATCT
ACAGCCAGCGCGGAGCCACCGGGCTCAAGAGGGTGGCTAAAGGACAGTCACAGAGGCTTCTTAGCCTGCTCTTCGCCAG
TGACAGATTATGTGAAGGTGGCAAAGCCACCACAAGCTAGACCACTGCTAAGAATAAGAGTGACTTTTAGAGGATGTTA
ATTGAAGGTTACAGCAAAATCGCCTGCTTTCTATTTCTCTATCTCAGAGTTCTGGTGCCACCTAGTGGTCAGAAGTAGA
ACTTGAAGCCCCAAGGTTTACTCAAAGGGCCAAAGGCATCATCAACGTTGTGAGAATTATCTTCTTCTGGCCTACCACA
GGACACCTCTGGGTTCTTCTCTGTGGTTACCAGGAAGCACAGTACTTACTAAATTTATGCTAACCATGACAAAAAGATTGT
CAACTCAAATTTGCTAGGAGTATTCTTTAGGTGCTGTCTGCAATTTTTTCTCTGTAAGTGAATGAAAAAGAAAAACAA
TAAAAATAAATTTGACTTCGAAAAA

SEQ ID. No: 51

FIGURE 24

>SC|19|14042 Frame 2
MAASRDADEIHKDVQNYGNNVLTTSADLQTNACVTRAKP
VPSYIRESLQNVHEDVSSRYYGGLTVPERLENCRIIDLGS GSGRDCYVLSQLVGEKGHV
TGIDMTEVQVEVAKTYLEHHMEHFGFQAPNVTF LHGRIEKLA EAGIQSESYGIVISNCVI
NLVPDKQQVLQEVYRVLKHGGEIYFSDVYASLEVPEDIKSHKVLWGECLGGALYWKDLAI
IAQKIGFCPPRLVTADIITVENELEGVLGDCRFVSATFRLFKLPKTEPAERCRRVYNGG
IKGHEKELIFDANFTFKEGEAVVDEETA AVLKNSRFAPDFLFTPV DASLEAPQGPF*

SEQ ID. No: 52

FIGURE 24

NACCTCTGTCAGAGCCTCGCTAGGAAACCTAGGAGTTTATAGAGGCCACTTGGCACCAGGAGCTAGCCGGGTAGGCGGAGCCT
 CACCTGGATTGAGTTCACAGCTGCCTAGACAGGCTCAGACTAGGTGCTGGGCACCTGGGAGGAGGAGGAGACTTAGCAG
 CAAAGGCTGTTAAACAGAGTGCCTGCCTAGGCTTGGAGGCAAGACGCTGCTGTTACAGTGGCAGAGCGAGGTATAGGAGTA
 TAATGGCTGTCCAGAGTGTGCGGCAGATGGCTCTACTGCTCTACTGAGTGTGTTTTCTCTGGTGCAAGGTGCACACAGTGGC
 AGCCCCCGAGAAGACTTCCGCTTCTGTGGCCAGCGGAAACAGACCCCAAGCAGCACCTCCCATATGATCAATCTTTCAGA
 GCCTCACATCTTTGTGGGAACACAGAGGAGACCCTCACAAATTCGTGCCCCCTTCTGGCAGCCCGAGATTTCCCGGCT
 TCTTCCACAGCCTAGAGGCTCTATCACTCTGCTGCTCTACTGAGTGCACACACTGGGAGACTCCACTTGGCTATGGC
 AAGCATGACTACCTGCTTAGTAGCCACCTCCAGACTCTCTGCTTCCAGAAACAGGAGCAGAGCTGAAGCAGGGGAGC
 CCGCTGATCGCCACCTCTGCAGTCTCTGGCAGATTCCCAGAACACCCAGCTGCTGGGGCTCCGAGGCTTCATCTTCT
 CTTCCACAATGCCCACACAAGGTCTCCCACAATGTCATCTGTGGACATGTGTGATCTCAAGAAGGAATTGCAGCAGCTT
 AGCAGGTACCTGCAGCACCTCTCAAAGGCTGCCAAGCGGCCACCCGAGCTTATCAGCAGGACGATCTACAGAGCCTGGA
 TCCAAAGCTGACCTCTGTGAGCTTCTCTGGGAGACACATTATCCTTTGAGGAGGACCGGGTCAATGCTACAGTGTGAAGT
 GTCCACCCACAGCCGGCTAGAGGATTCGATATCCACTCCCAAGAGGAGGAGGAGCAGAGTGAGGTCCAGGCATACTCG
 CTGTTGCTTCCCCGGGCGCTATTCCAGCAGCAGGAGGCGCTCGCCGGGATGAGCGCAAGGGCTCCTGGTAGTAGACTT
 CAGCAGCCAAGCTTTGTTCCAGGACAAGAATTCTAGCCAAGTCTCGGTGAGAAGCTTCTGGGTATTGTCGTGCGAGCA
 CCAAAGTCCAAACCTCTCAGATCCGGTGGTACTCACCTCCAGCACCAGCCTCAGCCAAAAAATGTGACTCTGCAGTGC
 GTGTTCTGGGTTGAAGACCGGCATCAAGCAGCAGGAGGAGTGTGAGCAGTGCAGGCTGCGAGACAGTGAAGCAGAGAC
 ACAGACATCTGCTGTGCAACCACTGACTACTTTGCAGTGTGCTATGCTATCCACAGAGGTAGAAGCCACTCACA
 AACACTACCTCAGGCTCTGTCTACGTGGCTGTGTCTCTGCTGCTGCTGCTGCTTCTTCACTATCGCTGCCACTCTC

>SC|28LONG|14041 Frame 2
MAVQVLRQMVFYFL
LSLFSLVQGAHSGSPREDFRFCGQRNQSTLHYDQSSEPHIFVWNTTEETLTIRAPFLA
APDI PRFFPEPRGLYHFCLYWSRHTGRLHLRYGKH DYLLSSQASRLLCFQKQE QSLKQGA
PLIATSVSSWQIPQNTSLPGAPSFIFSFHNAPHKVSHNASVDMCDLKKELQQLSRYLQHP
QKAAKRPTAAAFISQQLQSLESKLT SVSFLGDTLSFEEDRVNATVWKLPPTAGLEDLHIHS
QKEEEQSEVQAYSLLLPRAVFQQTRGRRRDDAKRLLVVDFSSQALFQDKNSSQVLGEKVL
GIVVQNTKVTNLSDFVVLT FQHQPQPKNVTLCVFWVEDPASSTGSWSSAGCETVSRDT
QTSCLCNHLYFAVLMVSSSTEVEATHKHYLTLLSYVGCVISALACVETIAAYLCSRRKSR
DYTIKVHNMNLLSAVFLLDVSFLLSEPVALTGSEACRTSAMFLHFSLLACL SWMGLEGYN
LYRLVVEVFGTYVPGYLLKLSIVGWGFPVFLVTLVALVDVNNYGP IILAVRRTPERVYTP
SMCWIRDSLVS YVTNLGLFSLVFLFNLA MLATMVVQILRLRPHSQNWPHVLTLLGLSLVL
GLPWALVFFSFASGT FQLVILYLF SIITSYQGFLIFLWYWSMRFAQGGPSP LKNNSDSA
KLPISSGSTSSRI*

SEQ ID. No: 54

FIGURE 24

CTCAGGGAAGGCCGATCTTCCGGGTGGAGGGGGAAGCGGCGTGACTGGAGTGGAAATTTTCCCAACACAACTTCTCGGA
GGCAACATATTGGAAGGGACTCGGGGAGGCCGGAGTCCAAATGGAAGTGGCTGAAAGAACTTCTCGCCCTGCTGATTCT
GAGCCCCGCGTCTGTCGCGCGGCCCTCAATTACCTCATCGACCTGTGGTCTTGACAGAACATTCTTCACAATCCAAAAGA
AAAAGCAGACTGGTTCGGGATGTTTGACGGCTATGATAGCTGCAGTGAGGACACAAGTAGGAGCTCCAGCTCTGAGGAGA
GTGAAGAAGAAAGTTGCTCCTTTACCTTCCAATCTCCAATCATCAAGAATAATGGACAAGTCTACACATACCCAGATGGT
AAATCTGGCATGGCTACCTGTGAGATGTGTGGGATGGTGGGTGTGCGAGATGCTTTTACTCTAAAACGAAACGTTTCTG
CAGTGTCTTCTGTTCAAGAAGTTACTCGTCAAACCTAAGAAGGCAAGCATTCTGGCCAGACTTCAGGGTAAGCCTCCAA
CAAAGAAAGCAAAAGTCTTCAAAAACACCTTTAGTTGCTAAATAGTGCCTATGCCAGTATCAAGCTACCTTGCAA
AATCAAGCAAAGACAAAAGCAGGCAATTTCTGCAATCTCTGTGGAAGGCTTCAGCTGGGGTAACATCAACAGCAACAG
CTTCATAGCAGCTCCAGTGGCCTGTTTTAAGCATGCACCTATGGGGACCTGCTGGGGTGATATCTCGGAAAATGTAAGGA
TAGAAGTTCCCAATACAGACTGCAGTCTACCTACCAAGTCTTCTGGATTGCTGGAAATTATAAAATTAGCAGGTTATAAT
GCCCTTTTGAGATATGAAGGATTGAAAATGATTCTTCTCTGGACTTCTGGTGCAATATATGTGGTCTGATATTCAATCC
AGTTGGTTGGTGTGCAGTAGTGGAAAACCTCTCGTCTCTAGAACTGTTCAACATAAATATACAACTGGAAAGCTT
TTCTAGTAAAAAGACTTACTGGTGCCAAAACACTTCTCTGATTCTCACAAAAGGTTTCTGAGAGTATGCAATATCCT
TTCAAACCTTGATGAGAGTAGAAGTAGTTGACAAGAGGCATTATGTGCAACAAAGAGTAGCAGTGGTGAAAGTGTAAAT
TGGAGGACGACTACGGCTGGTGTATGAAGAGAGTGAAGATGGAACAGACGACTTCTGGTGCCACATGCACAGCCCTTAA
TCCACCATATTGGATGGTCAAGAAGCATAGGCCATCGATTCAAGAGATCAGATATTACGAAGAAACAGGACGGACATTTC
GATACACCTCCACACTTATTTGCTAAGGTAAAGAAGTAGACCCAGAGTGGAGAATGGTTCAAAGAAGGAATGAAATTGGA
AGGTATAGACCCATTAAATCTTCTACAATATGTGTTGCCACCATAGAAAGGTGCTGGCTGATGGATTCTGATGATTG
GGATTGATGGCTCAGAAGCAGCAGATGGATCTGACTGGTCTGTTATCATGCAACCTCTCCTTCCATTTTCCCTGTGGGT
TTCTGTGAAATTAACATGATAGAACTGACTCCACCAGAGGTTACACAAAACCTTCTTTAAATGGTTTGACTACCTCAG
GGAACCCGGCTCCATTGCAGCACCAGTAAACTATTTAATAAGGATGTTCCAAACCACGGATTCCGTGTAGGAATGAAAT
TAGAAGCTGTAGATCTCATGGAGCCACGGTTAATATGTGTAGCCACAGTTACTCGAATTATTCACCGTCTCTTGAGGATA
CATTTTGTATGGTTGGGAAGAAGAGTATGACCAGTGGGTAGACTGTGAGTCCCGGACCTCAATCCTGTAGGGTGGTGTCA
GTTAACTGGATATCAACTACAGCCTCCAGCATCAGTCTCAAGAGAAAGCCAATCAGCTTCTTCAAACAGAAAGAAA
AGGCTAAGTCTCAGCAATACAAAGGACATAAGAAAATGACCACGTACAGCTGAAGGAGGAGCTGCTGGACGGGGAGGAC
TATAGCTTCTCCATGGAGCATCTGACCAGGAAAGCAATGGCTCTGCCACCGTCTACATCAAACAAGAGCCATGAGGCGA
CTCGGAAACCATTGGGACGGCGGGCTGTTACAGGACTGATTGGGAATCAGCCAGCTGTATAGCGGGCTATTCTACTGGG
ACATTTTGCTAAACACAGAAAAAGTTAGTTCCAGATTTTTCAGGTGGGGGGGAACTATTTTGGTGGGGGGCAATTT
TTCAATTTATAAGACGGACAATTTTGTGTTGATTGTAAGCTTTTGAAGAAATTTTGAATATTTTCCAAGTTTGGAT
TTATGTGCATTGTTAAACAAGAACTGAAATTATACTTTTGGTAAGATAAAAGTTTAGGTAGCAGGATTGAAGGAAATG
ATTAAGAAGGATATAGTTGTAATGSCACATGAACTGTCAATACAAATGAACCTTCTTGGTACCTGTTGGGAGATTTTGG
GATTTTAGAGTTAGGCCAGTCACATCTCCAGCTTCTTGTCTGCGAAAAATATGCAACTGAACCCCTCAGAGGGCTCAT
CACCATGTAGATCACGGAAGGGGTCAATTAATGTGCTCGCTGACGTTTATTGACAGCCATTTAACTGTTGTACAGAAA
CTTTTTCATTCTGCTAAAATTTATTTGGAGTTGTATATGAACTAGGAGAACTCTGGATACTTTTATATGCTCTCCTTCA
CGTAAAGATGATTAATAATTTGCTAACACTAAAGTTTAACTGGAATGGTTGACAAAGTACACAAGATCTCAGTCTACAT
AAAGGGTTGGGGGAAATTAGTATTTTCTAAGTTTATCTGTTTTCTTGTAGAAAAATCCCATAAAGGGATATTCTT
GTCAGACCTCTGCCATTTCTTCCATCTGTGAGAGAGCAACGTGATGTAACCAACCTGAGAGCAGGTACTCTCTGTGTT
ACAAAAGAATCACATGGAATTTGTTAACCAAAATAAGTGTAGATTTCAGAATCATTCCAGCTCTACTCTTAACTGTCTT
GAATTTGTTAGAACTGATTTGAAGAAGATGTTTCTTAAATTCATACACACACACATATTGAATTTGTCTTTA
ATTGTAGTCAGAAATTTATTATAACAATATTGATTGGACCATTTTAAACATTCCTTATTTTAAATTCATACGGCTTCC
TTAAGAAGTAGAATGAAGGGTAAATTTGGTGACATGTTTGTCTCTGCAATTTTCTAACCTTCTACCGAATTTGTGACTGAC
TCAGAGAGCTCTAGCATTTACCAGTGAGGTTGAGAACTAACTCTCAGGAATTCATTGCATCTGTTGTAGAAGTGCTTC
TGGGTCTGAGCCTGGCCTCCTCAGAGTGGTAATACTGCCACTTCTCTGGAATAAGGCAGGGCTAATGAGAACTAAT
CAAATGATTAACCTCTGCGGCCCTCAGCCTTTGGAATGCTAAAAA

SEQ ID. No: 55

FIGURE 24

SCisa49p01-full.seq|14351 Frame 2
MFDGYDSCSEDTSSSSSSSEEEVAPLPSNLPI
IKNNGQVYTYPDGKSGMATCEMCGMVGVDAFYSKTKRFCVSCSRSYSSNSKKASILAR
LQGKPPTKKAKVLQKQPLVAKLAAYAQYQATLQNOAKTKAGNSAISVEGFSWGNYSNS
FIAAPVACFKHAPMGTCWGDISENVRIEVPNTDCSLPTKVFWIAGIIKLAGYNALLRYEG
FENDSSLDFWCNICGSDIHPVGWCAASGKPLVPPRTVQHKYTNWKAFLVKRLTGAKTLPP
DFSQKVSESMQYPFKPCMRVEVVDKRHLCTRVAVSVIGGRLRLVYEESEDGTDDEWC
HMHSPLIHHIGWSRSIGHRFKRSDITKKQDGHFDTPPHLFAKVKEVDQSGEWFKEGMKLE
AIDPLNLSTICVATIRKVLADGFLMIGIDGSEAADGSDWFCYHATSPSIFPVGFCEINMI
ELTPPRGYTKLPFKWFDYLRGTGSIAAPVKLENKDVPNHGFVRVGMKLEAVDLMEPRICV
ATVTRIHRLLRIHFDGWEEYDQWVDCESPDLNPVGWCQLTGYQLQPPASQSSRESQSA
SSKQKKKAKSQYKGHKKMTTSQLKEELLDGEDYSFLHGASDQESNGSATVYIKQEP*

SEQ ID. No: 56

FIGURE 24

>SC|sa61|14347
GGGGCTTCACTCTCCAGCCTTGGAGAGCAGTCAGGGAAGGCCTCCAGGAGGAAGCAGCTACCTGTGGCAGGGCAGAGCAG
TGCCTGTACCAGTCTCCCGAACTTCAGAACTGCGGTGGCCATGGAGTCTGTGGAGCCCCCGGTCAAAGACGGCATCCTC
TACCAGCAGCACGTAAGTTTGGCAAGAAATGCTGGCGCAAAGTGTGGGCTCTGCTGTATGCGGGAGGCCCATCAGGGGT
AGCTCGGCTAGAAAGCTGGGACGTGCGTGATGGTGGCCTGGGACCAGCAGGCGACAGGTCCACAGGGCCCCAGCCGTCGAG
GGGAACGCCGGGTACATACGCTTGGCTGACTGTGTATCTGTCTGCTGCGGATGGCGAGAGCTGTCCCAGGGACACTGGT
GCCTTCCTGATTACCACCACTGAGCGAAGCCACCTGTTGGCTGCACAGCACCGCCAGTCCTGGGTGGACCCCATCTGTCA
GCTGGCCTTCCCGGTACCGGAGAATGTTCTGTCAGGATCAGGACAGGCTGAGAATCCAAAAGGGGCTTTGTTCCCATGG
AAGAAAACCTATCTACTCCTCCTGGCAGGAAGTGACCGAATTTCCGGTGATCGTGCAGAAGACAGAAGCCACCTCCCGC
TGCCAGCTGAAAGGACCCTACCTCCTGGTGCTGGGCCAAGATGACATCCAAGTGAAGGAGACATCCAAGCCCCAGGCCTG
TTTGTAGCTGGCCCTACCGTTTCTGCGCAAGTACGGCTCTGACAAGGGTGTGTTCTCGTTTGAGGCTGGCCGCGCTGTG-
ACTCAGGTGAGGGCCTTTTTCCTTCACTAGCCCGCTGCCCCAGACATATGTGGGGTTGTGGCTGCCGCCATTGCCCGC
CAGCGGGAGCGTCTTCCAGAGCTGGCCATGTCCCCACCTGCCCTCGGGCCCTCTCCCTGCCCTCCCTAGAGCC
CCCTGGAGAGCTTCGGGAGGTGGCCCCAGGATTGAGCTGCCCACTCCAGAAAGCTGCCCTCTAACTGATCCCGGCCCTC
AAAGCCTACCATTTGCTGCTCAGCCCCAACACAAGAAGGACCGGCATCCGGTCTCTATGCGTCCGTGTGCAAGCAGACCAG
CAAGCACACAGGCACGGCGGAGCATTTCTATGAGAAGCTGTGTCATGCTGGAGGCCAGCCTTGGGCTGACCAATGGGGTCT
CTGAAGCCCCAAGAGGGCCCCCTTGGTGGCCGACGCCCTTGGGCAGCCTTATCTACCATAACACTGAGGATCTGAGTTG
GCCGGGCTCGGCCAGGACAGCAATCTGGAAGCCAGTACCGGAGGCTGCTGGAAGTGGAGCTGGATGAGGCCGGAAGCG
CCGGCCGCTCTGGAGCGCAGGCAGGCATCAAGGCCAAGCTGGTGACCTGCTGACCCGTGAACGGAAGAGGGCCCCGCC
CCCTGTGACCGGCCCTGAAGGCCTGAGCGGCCAGCCACTGCAGGACAGAGGTGATCACCCAAGACCAGGAACAACCTTCGA
ACATAACCCGCTCTACTCTGACCTGCAGGGACAAGCCAGGTGGCCCCGGGAGGAGCCACTCTGCCCTACCTCCTCCCTC
AGACTGTACAGATTGAACAGTAATAAAGCTTGCCCTATCAACTTCAAAAAA

SEQ ID. No: 57

FIGURE 24

89/99

SC|sa61|14347 Frame 2
MESVEPPVKDGILYQQHVKE
GKKCWRKVVALLYAGGPSGVARLLSWDVRDGGGLPAGDRSTGPSRRGERRVIRLADCVSV
LPADGESCPDRTGAFLITTTERSHLAAQHRQSWVDPICQLAFPGTGECSGSGQAENPK
RGFVPMEEINSIYSSWQEVTEFPVIQKTEATSRCLKGPYLLVLGQDDIQLRETSKPOAC
FSWPYRFLRKYGSDKGVFSFEAGRCDSGEGLFAFSSPRAPDICGVAAAIARQRERLPE
LAMSPPCPLPRALSLSLEPPGELEVAPGFELPTPRKLPLTDPGPQSLPILLSPNTTRT
GIRSLCVRVQADQQAHRHGGAFLRVHAGGQFWADQWGS*

SEQ ID. No: 58

FIGURE 24

90/99

>SC|smc34|14348
GGCCTCCGACATATTGCCCGCAGGAGCTGCGGCGGTGAGCGGAGAGCGCCGGGGGAAGGAGATGGGAGGACGAAGAGGTC
CCAACAGGACATCTTACTATCGAAACCCGCTCTGTGAGCCAGGATCATCAGGGGCCCTCTGGTGGAGGCCACTCTTCTAGT
GCATCCGTGAGCAGCGTCCGTTCTCGAAGCAGGACCCTTCTGGGACGGGCCCTCTCCAGCCCTCCGCTGGCCGCGCAGAC
CGTTGTGCCTCTACAGCACTGCAAGATCCCCGAGCTGCCCGTCCAGGCCAGCATCTGTTCGAGCTACAGCTCTTCTTCT
GTCAGCTCATAGCCCTGTTCTGTGCACTACATCAACATCTATAAGACAGTCTGGTGGTATCCACCCTCGCACCCACCCTCC
CACACTTCCCTGAACTTCCACCTGATCGACTTCAACTTGCTGATGGTGACCGCCATTGTTCTGGGCCGCGGATTTCATCGG
GTCCATCGTGAAGGAGGCTTCTCAGAGGGGGAAAGGTCTCCCTCTTCCGCTCCATCCTGCTGTTCTCACCCTGCTCAGG
TTCTCACGGCGCAGGCTGGAGTCTGTGCCGGTCCCTCATCCACCTCTTCAGGACCTACTCCTTCTGAACCTCCTGTTT
CTCTGCTATCCGTTTGGGATGTACATTCCGTTCTCTGCAGCTGAACTATGATCTTCGCAAGACGAACCTCTTCACCCACAT
GGCTTCCATGGGACCCCGAGAGGCAGTCAGTGGCCTGGCAAGGAGTCGGGACTACTTTCTGACACTGCGGGAGACTTGGG
AGCAGCATACACGACAGCTATATGGCCCGGAAGCCATGCCACCCATGCCCTGCTGCTTGTACCCAGCCCTCATTCGCAAT
GAAGTTGAGTTCCTCAAAATGGACTTCAATTGGCGAATGAAGGAAGTACTTGTGAGCTCCATGCTGAGTGCTTACTATGT
GGCCTTTGTACCTGTGTGTTTGTGAAGAATACACATTACTATGACAAGCGCTGGTCTGTGAGCTCTTCTGCTGGTGT
CCATCAGTACCTCCGTGATCCTCATGCAGCACTTGTGCTGCCAGCTACTGTGACCTGCTGCACAAGGCCGCGCCAC
CTGGCGTGTGGCAGAAGTGGACCCAGCTCTGTGCTCCAATGTGCTGCAGCCACCGTGGACTGAAGAGTGCATGTGGCC
ACAGGGTGTGCTGGTGAACACAGCAAGAATGTCTACAAAGCAGTGGGCCACTACAACGTGGCATCCCTCCGATGTCTC
CCACTTCCGCTTCCATTTCTTTTTCAGCAACCCCTCGGGATCCTCAACATCCTTTTGTGCTGGAGGGTGCCGTGCTC
TCTACCAGCTGTACTCCTTAATGTCTCAGAGAAGTGGCACCAGACCATCTCCCTGGCCCTCATCCTCTTCAGCAACTAC
TATGCCTTCTTCAAGCTGCTCCGGGACCGCCTGGTATTGGGCAAAGCTACTCCTACTCAGCCAGCCCCCAGAGGACCTGG
ACCACCGGTTCTCCTGAGTCCCGGGGTGATCCTCCAAACAACATCCGGGCTTCACCCAGGGGCTTCTGGCCAAGGGCTC
TTGGGGTGGGGGGAGGGATACAAAAACAACAAAAACAACAAAAAACCAACCCACCCAGAGCTTTGTATTTT
TGTTACGTAAGTCTTCTTTGATAATTGATGTGTAAGAAAAAGTCTTATTTTATACTCCAAAA

SEQ ID. No: 59

FIGURE 24

91/99

SCismc34114348 Frame 2
MGRRGPNRTSYRNP LCEPGSSGASGGHSSASVSSVR
SRSRTTSGTGLSSPPLAAQTVP LQHCKIPELPVQASILFELQLFFCQLIALFVHYINIY
KTVWWYPPSHPPSHTSLNFHLIDFNLLMVTAILGRRFIGSIVKEASQRGKGLPLPLHPA
VPHPLHVLTA TGWSLCRS LIHLERTYSFLNLLFLCYPFGMYIPFLQLNYDLRKTNLFTHM
ASMGPREAVSGLARSRDYFTLRETWKQHTRQLYGPEAMPTHACCLSPSLIRNEVEFLKM
DENWRMKEVLVSSMLSAYYVAFVPVWFVNTHYYDKRWSCELFLLVSISTSVILMQHLLP
ASYCDLLHKA AHLACWQKVDALCSNVLQFPWTEECMWFPQGVLVKHSKNVYKAVGHYNV
ASPPMSPTSASISFSATPCGSSTSFCCWRVPSLSTSTP*

SEQ ID. No: 60

FIGURE 24

92/99

>SC1AA4.1|15737
GTGAAAGCAGCAGTGGCGCTCTGCTCCCTTCAGAGCACAGCCTGGTGTCAAGGTCCAGGTTCCACCGGCTGCTGCTGTCA
CCGCAGGGGAGTCTAGCCCCCTCCAGAGGAGACAGAGAATGGCCATCTCAACTGGTTTGTTCCTGCTGCTGGGGCT
CCTTGGCCAGCCCTGGGCAGGGGCTGCTGCTGATTACAGGCTGTGGTGTGCGAGGGGACTGCCTGCTATACAGCCCAT
GGGGCAAGCTGAGTGCCGCTGAAGCCAGCATCGCTGCAATGAGAATGGAGGCAATCTTGCCACCGTGAAGAGTGAGGAG
GAGGCCCGGCATGTTCAAGCAAGCCCTGACTCAGCTCCTGAAGACCAAGGCACCCTTGAAGCAAAGATGGGCAAAATCTG
GATCGGGCTCCAGCGAGAGAAGGGCAACTGTACGTACCATGATTGGCCAATGAGGGGCTTCAGCTGGGTGGGTGGTGGAG
AGGACACAGCTTATTCAAACTGGTACAAAGCCAGCAAGAGCTCCTGTATCTTTAAACGCTGTGTGCCCTCATACTGGAC
CTGTCTTGACACCTCACCCAGCCATCTGCCAAGTGGCATGAGAGTCCCTGTGGGACCCCGAAGCTCCAGGTAACAG
CATTGAAGGTTTCTGTGCAAGTTCAACTTCAAAGGCATGTGTAGGCCACTGGCGCTGGGTGGTCCAGGGCGGGTGACCT
ATACCAACCCCTTCCAGGCCACTACCTCCTCTCTGGAGGCTGTGCCTTTTGCCTCTGTAGCCAATGTAGCTTGTGGGGAT
GAAGCTAAGAGTGAAACCCACTATTTCTATGCAATGAAAAGACTCCAGGAATATTTCACTGGGGCAGCTCAGGCCCACT
CTGTGTGAGCCCCAAGTTTGGTTGCAGTTTCAACAACGGGGGCTGCCAGCAGGATTGCTTCGAAGGTGGCGATGGCTCCT
TCCGCTGCGGGCTGCGGGCTGGATTTGAGTCTGCTGGATGATCTAGTAACCTTGTGCCTCCAGGAACCCCTGCAGCTCAAAC
CCATGCCACAGGAGGTGGCATGTGCCATTCTGTACCACTCAGTGAAAACCTACACTTGCCGTTGTCCAGCGGGCTACCAGT
GGACTCTAGCCAAGTGCACTGTGTGGATATAGATGAGTGCCAGGACTCCCCCTGTGCCCAGGATTGTGTCAACACTCTAG
GGAGCTTCCACTGTGAATGTTGGGTTGGTTACCAACCCAGTGGCCCCAAGGAAGAGGCGCTGTGAAGATGTGGATGAGTGT
GCAGCTGCCAACTCGCCCTGTGCCCAAGGCTGCATCAACACTGATGGCTCTTTCTACTGCTCCTGTAAAGAGGGCTATAT
TGTGTCTGGGGAAGACAGTACCCAGTGTGAGGATATAGATGAGTGTTCGGACGCAAGGGGCAATCCATGTGATTCCTGT
GCTTCAACACAGATGGTTCTTCAGGTGTGGCTGCCCGCCAGGCTGGGAGCTGGCTCCCAATGGGGTCTTTTGTAGCAGG
GGCACTGTGTTTTCTGAACCTACCAGCCAGGCTCCCCAAAAGGAAGACAACGATGACAGAAAGGAGAGTACTATGCCTCC
TACTGAAATGCCAGTTCTCCTAGTGGCTCTAAGGATGTCTCCAACAGAGCACAGACAACAGGTCTCTTCGTCCAATCAG
ATATTCCTCACTGCCTCTGTTCCACTAGAAATAGAAATCCCTAGTGAAGTATCTGATGTCTGGTTCGAGTTGGGCACATAC
CTCCCCACGACCTCCGGCCACAGCAAGCCGACACATGAAGATTCTGTGTCTGCACACAGTGACACCGATGGGCAGAACCT
GCTTCTGTTTTACATCCTGGGGACGGTGGTGGCCATCTCACTCTTGCTGGTGGTGGCCCTAGGGATTCTCATTTATCATA
AACGGAGAGCCAAGAAGGAGGAGATAAAAGAGAAGAAGCCTCAGAATGCAGCCGACAGCTATTCTGGGTTCCAGAGCGA
GCAGAGAGCCCAAGCCCCGGAGAATCAGTACAGCCCAACACCAGGACAGACTGCTGAAGACTATGTGGCCTTAGAGACAG
CTGCCACTACCTTCAGAGCTACCTTCTTAGATGAGGGGAAGCCACATCATTCTGAATGACTTGACTGGACTCTCAGCAA
AAAAATTGTGCACCTTCCACTTAAGAACCTGGTGGCTTGGGATAGGCAGGTATTTCTTGGTGCCTTTGATATGTCTGGG
GGTGAAGCTGTGTGTGGTTGTGCTTGTGGGGAGTTTGTGGATATTGACAGACCTCACTCAAACACCCCTTTCAAAT
CCAATAGCAACTGGTTCCTCTGGTTCCTAATTAGGGGGAAGGAGTCAAGGGGTGGGACAGGGTGGGGGGATGGGGCTT
CAAAGTTTTTTCTTATCACTTGATTATCATCGAAGGAGTTACTGGTGCTAATTACAATGGAAACAGTTCTCTTCCATCA
CAGGACAGACACACCTCAATCCTCCATGGGGTCAACAACATATATACCCCAAGTGACCCCTTAGGCAAGGACTTGTGAGA
ACTGCATCACATTTTACCTGTTCTCAACAGTACCCATCTATTTCAAGTGGGATCTCTGGACCTTTCTCTCTCCCATCT
TGTCTGCAATGTGGCAAATGGCTTCTTTTGCATTTTACTCCGCCCCACCCCAAGCTGAAGTTCAATTTGCAGATCAGC
GATTAAGTCTGAATTGTGGTGGTCACTCTTGTTCCTTTTGCAGGGGTTATTGTAAATGTTAGTAATTTGCGCTCAA
GCCCTCAGTAAGAACAATAATATTTAAATATGTGCGTTTGAATCTGTTTATGCATCCTGGAACCTGTGGGATGCTCA
GGCAAGAGTGACTTTAGTCTTTCAGTGAATGTTGCCAGAATGTGGGTAGGGAAGGCTCACAGGTTACTCTCTCTCTTAG
AGCTACAACATAACATTCTGAGGGGAGTCACAGGGTTCCTTTAAAAAGTGGGAGCTATGTCATGCTTTGAGCTTTCTGT
TAAGCACCTCTCCTAATAAATCTGAAAAA

SEQ ID: No: 61

FIGURE 24

93/99

SCIAA4.1|15737 Frame 3
MAISTGLFLLGLLGPWAG
AAADSQAVVCEGTACYTAHWGKLSAAEAQHRCNENGGNLATVKSEEEARHVQOALTQLLK
TKAPLEAKMGKFWIGLQREKGNCTYHDLPMRGFSWVGGGEDTAYSNWYKASKSSCIFKRC
VSLILDLSLTPHPSHLPKWHESPCGTPEAPGNSIEGFLCKENFKGMCRLALGGPGRVTY
TTPFQATTSSLEAVPFASVANVACGDEAKSETHYFLCNEKTPGIFHWGSSGPLCVSPKFG
CSFNNGGCQQDCFEGGDGSFRCGCRPGFRLLDDLVTCASRNPCSSNPCTGGGMCHSVPLS

ENYTCRCPSGYQLDSSQVHCVDIDECQDSPCAQDCVNTLGSFHCECWVGYQPSGPKKEEAC
EDVDECAAANSPCAQGCINTDGSFYCSCKEGYIVSGEDSTQCEDIDECSDARGNFCDSL
FNTDGSFRCGCPGWELAPNGVFCRGTVESELPARPPQKEDNDRKESTMPPTMPSSP
SGSKDVSNRAQTTGLFVQSDIPTASVPLEIEIPSEVSDVWFELGTYLPTTSGHSPKPTHE
SVSAHSDTDGQNLLEYIILGTVAISLLLVLAGILYHKRRAKKEEIKEKKPQNAADSY
SWVPERAESQAPENQYSPTPGTDC*

SEQ ID. No: 62

FIGURE 24

94/99

SC|HDD2|16960
CCCATTACTTCCACCTCCCATGAGGACTGCACCTGCGGCAGCGATTTTATGTGAACTTGGATTACTGTATGGAAATGGAG -
GAGTGAGAAAGTGTGGGAATCATAAAAGAAGGCAGAACCAAAGACAGAGGGATCTTGGGCCATGAGTCACAGCAAGGAA
AGCCTCCCACCAAACATCTGCACTGGACTCCTACAGGAACAAGAAATGAACTATCGTGTTCAGCCCCCTGAGACTTGGGCT
TAACATGATTTTACATTGTCTACCTACCTAATAGAGCAGAGATGTAATATTATTCTTATTTAGAGGTGTGATGCCT
CAGCTGCAATGGGTGAGAACTACTCCTCATTTATTCTCTTCCAAGGCAATAAAAGAGAATGGACCAAAGACAGTCTGTCA
TCACATCTAGTCAAAAGAGCTAATGTCGCAGTACAACCTTTCAAAGAAAAAGAAAAACAAGAAAAAGTAATAACAG
ATGTGTTCTGCTTGAAAAA

SEQ ID. No: 63

FIGURE 24

95/99

>SCIHDD2116960 Frame 2
MRTAPAAAILCELGLLYGNGGVRKCGNHHKRRQNQKTEGSWAMSHSKESLPPNIC
TGLLQEDEMNYRVQPLRLGLNMIFTLSTLPNRAEM*

SEQ ID. No: 64

FIGURE 24

96/99

ACGGAGACAGCTGCAGCGGGGACAGTATTCAGGACAGCTACACCGGCATGGAAAACCTCCGACAAGGACGCCATGAACAGC
CAATTTGCTAATGAAGATGCCGAAAGTCAGAAAGTTCCTGACAAATGGGTTTTAGGGAAAGAAAGCTAGCCGATTACGC
GGATGAGCATCACCTGGAATGACTTCCTTTGGAATGTCCTCATTTAACCTGAGCAACGCCATCATGGGCAGTGGGATCT
TAGGCTTGCTATGCCATGGCCAAACCCGGGATCATCTTTTATAATCATGCTGCTTACTGTGGCAATACTCTCGCTC
TACTCGGTTACCTTTTGTGAAGACAGCCAAGGAAGGAGGGTCTCTAATCTATGAAAAATGGGCGAGAAAGCATTGGG
ATGGCCTGGGAAAATTGGAGCCTTCATCTCTATTACAATGCAGAACATTGGAGCCATGTCAAGCTACCTCTTCATCATT
AGTACGAACTGCCTGAAGTAATCAGAGCATTTCATGGGACTTGAAGAAAACACTGGGGAATGGTACCTCAACGGCAACTAC
CTCGCTTTATTTGTGTCGTGGGGATCATCTCCCGCTCTCTCTCTTAAAAATTTAGGCTACCTTGGCTACACCACTGG
ATTTTCTCTCTCTGCTGATGGTGTCTTTCGTGAGTGGTGTATTTACAAAAATTTCAAATTCCTGCCCTCTGCTGCTC
TGGATCACAACAACGAAATCTGACGTTCAACAACACACTTCGATTACATGATCTCGCTGCCTAATGACTCGGAGAGC
TCGGGTGTGAACCTTCATGATGGATTACGCTCACCACAACCCAGCTGGGCTGGATGAGAAGCAGGTCCGAGGCCCTCTTCA
CAGCAATGGCGTGGAGTACGAAGCCCAAGGCTGCTGAGAAATGCCAACCAAAATACTTTGTGTTCAATTCCTGGAGCGGCT
ATGCAATCCCAATCTGGCTTTTGTCTGCTGCCACCTGAGGTCTTCCCATCTACAGCGAGCTTAAAGATCGATCC
CGCAGAAAGATGACAGCGGTGTCCAACATTTCCATCTCAGGATGCTCGTCTGATGACCTTCTTGGCGCCCTCTTTGGTTA
TCTGAGCTTCTACGGGGACGTTGAAGACGAGCTGCTGCATGCTTACAGCAAGGTCTACACATTTGATACGGCTCTTCTCA
TGGTGGCCTGGCAGTCTGGTGGCAGTGACACTGACCGTGGCCATCGTGTCTTCCGATCCGTACTTGGGTGATCACA
CTGCTGTTTCCAAGGAAACCTTCAGCTGGCTGAAGCATTTCCGGATCGCTGCAATCATCATCGCACTCAACAACATCT
GGTCACTCTCGTGCCTACCATCAAATACATCTTTGATTCTAGGGGCTTCTTCTGCCACTATGCTGATTTTCATTCTTC
CGGTGCGTTTATCTCAAGCTCGTCAAGAAAGAACCTCTAAGATCACCCAGAAAGATTGGGGCTTTGGTCTTCTTGTG
ACTGGAATTTTTCATGATGGGAAGCATGGCGCTCATTATCTCGACTGGATCTACAACCCGCCGAATCCCAATCACCA
CTAATCCCGGGGAGACCGCTCTCCACTGGAAACAGCTGAAATGTCTGAAGGACATTTAGTTGTCTTGATTGGGATGTT
AGTCTGAGGAATTAGCAAGATTCCAAAGACGTTTTCTAGCTCTATCAGCACACATTTAACCCAGGCGGTGCGAGTGCA
TGTGTGATGCCCGAGTTGTGTTTGCAGCAGCTGTGCAAGCTGAAGCCTGTTGGCTGCGTGTGTTGGTCAGCAGACAATAG
CCTGTCCCCCATGGTCACTCCACTTCTCTCACCCCCAGATTAACAGGTAATCTACTCTCAGAACATCAGACAAAGAC
CTCTGGTTGGGATACTTGTGGAAGAGAAAATTATGGGTTTTGTTGGGAATGGTTTTGTTGGGAATGGTGAAGGATGCAT
TAAAAATCTGTGCGAAGTATCATCAGTTACGGCCATCTCTCACTCTACCAACACTAAGGGTGGTTGACTAGCTGAG
GCAGGGGATATCTTGGGCTGTCCCTGTGAGGATCATGACGTATGACGGTGGCAGTATAGAGTACTTCAATTTCAATACT
CAAGGAATAGTTTGCCCAACCTGCTTATTACCCGAGTTAGTGAATTTTCTTCTTTTATTTTGGTATGGCAATGT
TCAAACATCTCACTAAACATAGAGGGGGTTATTAGTTTATCATGAAACACTACCAAAAGTAATGGTTTTTAAACATC
TGCTTTTCTATGTTGTTTCTGATATCTTTGTTCCACTTGTCTTTGAACAAGGCTTCCCTCTCGTCTTCCGTTCTCACT
CTTCGTAATTTTCAGAAAACACAGTCCGTGTGGGAGACACTACCCAGTATTGTTGATACATCTCTATTTGATAA

ACATTCACTGAGGAACTGTGATTTTGTATATGTTTGTGTACATAATCTCATTCTGCAGTTATCAGAAGCTTGACATA
TGGGACATTGGATTTTATTTTACATATGAGGTTTTTTCTTCCAGACAAAATGTTTATATCATCAGGGTGGGGC
AGGGAATTAAGCTGGTGGGCTCAAAAATCCATGGGTACTTATCTGTCCATTGGAGACATCTAAAAATTAAGTCAAAGTT
GTGCAATAGTTCACTAATGCTCTTCACTGTTTACAAGACTATAATCATCTCAGCAAGTAGCGAAGGAAGGGTTTGTAAAT
AATTTTCACTGCTAACAATGGTCTTTGAATTTAATCATGTCTAGAATAGGGTTTAGTTCATTTTAGTTTACATCGGAACTT
GGTGACATTCTTGGCATTAAATAGAACATCGTGTGGTTAGGAATAAATGCACCAACATGACCCAAAACCATTTGCTCA
CATCCATGGACGCTTACACGGGATGGAGTTGAGACTCCGCCCCACTGCTAGAATCTCTGGCATTGACTTCTTTTTTCAA
TTACTGGTTGTATATTTTACATTTCACTAAAAAGCCAGTCTGAAATGCAAGTATCTACAGGTAGTGATATACTGTCAC
ACCAAAGCAGTTTTTCAACAGTGTGCTAAGAAACATATATTTGTTTACTGTGTCATTTCTGGACCTTAAGACTATACA
ATGGTGTGCTTTGTGATAGAAATGAGCATTACTAGAAACTACCAATACTTGGAGTTATTATCATTTGGGAGGGAAAA
AATCACCAAATAGGATCTTAAAGATGTCTTTCTATTTTCATTTCATCCAAGAAATAAATCAAGATAGAAAGCAGAAAGGGT
GTGCTGCAATATTAATTTATTTTATCTGAACATATTAATGTGAATATAAAGATTAAATCCAAAAATATTGGAAACC
TGTGAAATCTGTACGTATTTCCGATCTTTAATCTTAGAGCAGCTAACCTGGAAGTATCCACTTGTGTAGTTAGATATTCA
AGAGCAATATCACAGTGCCAAAGACTTCATTCACACAGCACCAGTCAATTTGTGGCCGTTTTAGGATTAGAGTTTACA
CAGACTCCAGGGCAGATACTCAGAAATACCGTTTCTGGTAAATTTAACCCATTGTAAACAGATACAAATTTATTTCT
TACATCATTTATAAGACGGTTCAATGTACTGGGGGATTTTTTTTATTACAGACTGTGTATTGGTATATAATAATAAAT
TTTTCAATGAAAAAA

SEQ ID. No: 65

FIGURE 24

97/99

SC|aftpl115065 Frame 2
MDPMELNNVSI EPDGDSCSGDSIQDS/TGMENSDKDAMNSQFANEDA
ESQKFLTNGFLGKKKLADYADEHHPGATSGMSSFNLSNAIMGSGILGLSYAMANTGIIL
FIIMLLTVAILSLYSVHLLKTAKEGGSLIYEKLGKAFGWPGKIGAFISITMQNIGAMS
SYLFIKYEPEVIRAFMGLEENTGEW/LNGNYLVLFVSVGIILPLSLLKNLGYLGYTSG
FSLSCMVFFVSVVYKKFQIPCPALPDHNNGNLTFNNTLPIHMSLBNDESSGVNFMM
DYAHNPAGLDEKQVAGPLHSNGVEYEAQGAEKCPKYFVFNSTAYAIPILAFAFVCHP
EVLPIYSELKDRSRKMQTVSNISISGMLVMYLLAALFGYLSFYGDVEDELLHAYSKVYT
FDTALLMVRVLAVLVAVTLTVPIVLFPPIRTSVITLLFPKPFSWLKHFGIAAIIIALNNIL
VILVPTIKYIFGFIGASSATMLIFILPAAFYKLVKKEPLRSPQKIGALVFLVTGIIIFMM
GSMALIILDWIYNPPNPNHH*

SEQ ID. No: 66

FIGURE 24

98/99

>SC|F77|12685
AGCGACATGGCCCCGCGCTCCAGGCCAGCCTCCAGGCGGCTCTCAACTGAGGTTCTGCTGTTCTGCTGCTGTT
GCTGCTGCTGCTGTGTCATGGCCATCGCAGGGGGACGCCCTGGCAATGCCTGAACAGCGACCCCTCCGGCCCTGAGTCCCAAC
TCAACGCCGACGAGCTACGGGGTCGCTTCCAGGACCTGCTGAGCCGGCTGCATGCCAACCAGAGCCGAGAGGACTCGAAC
TCAGAACCAAGTCTTGACCCAGCTGTCCGGATACTCAGTCCAGAGGTGAGATTGGGGTCCCACGGCCAGCTGCTACTCCG
CGTCAACCGGGCGTCGCTGAGTCAGGGTCTCCCCGAAGCCTACCGCGTGCACCGAGCGCTGCTCCTGCTGACGCCGACGG
CCCCGCCCTGGGACATCACTAGGCCCTGAAGCGTGGCTCAGCCTCCGGGGACCCCGTGTCCCGCATTACGGCTGCGC
CTGACGCCGCTCCGGACCTGGCTATGCTGCCCTCTGGCGGCACGCAGCTGGAACTGCGCTTACGGGTAGCCGCCGCGAG
GGGGCGCCGAAGCGCGCATGCGCACCCAAGAGACTCGTGCCCACTGGGTCCGGGGCGCTGCTGTCACTTGGAGACTGTGC
AGGCAACTCTTGAAGACTTGGGCTGGAGCGACTGGGTGCTGTCCCCGCCAGCTGCAGCTGAGCATGTGCGTGGGCGAG
TGTCCCCACCTGTATCGCTCCGCGAACACGCATGCGCAGATCAAAGCACGCGCTGCATGGCCTGCAGCCTGACAAGGTGCC
TGCCCCGTGCTGTGTCCCTCCAGCTACACCCGGTGGTTCTTATGCACAGGACAGACAGTGGTGTGTCACTGCAGACTT
ATGATGACCTGGTGGCCCGGGGCTGCCACTGCGCTTGAGCACCGGGCCCTGCTCCTCACCTACACTCCCTTCAAGGATG
CTATTTATATTTGATTTATTAATATTATTAATTTATTGGGGTGGGTGGGTGGATGGATTGTGATTTATTTAAACT
CTGCTAATAAAGGTGAGCTTGGTTTCAAAA

SEQ ID. No: 67

FIGURE 24

>SCIF77/12685 Frame 1
MAPFALQAQPPGGSQLRFLFLLLLLLLLLSWPSQGDALAMPEQRPSGPESQLNADEL
GRFQDLSRLHANQSREDSNSEPSDPFAVRILSPEVRLGSHGQLLLRVNRASLSQGLPEA
YRVHRALLLTPTARPWDITRPLKRALSRLGPRAPALRLRTPPPDLAMLPSGGTQLELR
LRVAAGRGRRSAAHPRDSCPLGPGRCHLETVQATLEDLGWSDWVLSFRQLQLSMCVGE
CPHLYRSANTHAQIKARLHGLQPDKVPAPCCVPSSYTPVVLHRTDSGVS LQTYDDLVAR

SEQ ID. No: 68

FIGURE 24

SEQUENCE LISTING

<110> Lemischka, Ihor
Moore, Kateri

<120> GENES THAT REGULATE HEMATOPIETIC BLOOD FORMING STEM
CELLS AND USES THEREOF

<130> 2275-1-005PCT

<140> UNASSIGNED

<141> 1999-08-20

<150> US 09/138,132

<151> 1998-08-21

<160> 96

<170> PatentIn Ver. 2.0

<210> 1

<211> 400

<212> DNA

<213> Mus musculus

<400> 1

cccacgcgtc	cggagagcgg	cagccatctc	catgccagg	ttgaacaaaa	tgccaggaat	60
gttcttttct	gctagcccaa	aggattcgaa	agaacacagc	cattctcttc	tagacnacia	120
aaagcagaaa	aaaaggccaa	agacttttgg	aatggacgtg	aaaacatacc	tgagatcgat	180
gatcccatat	ctggaatctg	ggatgaaatc	ggccaagtcc	aaagacatac	tttctgctga	240
agaagtatat	cagtgtgtctc	agtctctgga	aaaatccttg	ccaaccagac	aggtcnaaat	300
gtctttggaa	gattctaaag	tctgattcag	tgaggaaaaa	attgaattct	ggttggcttg	360
tgaggactat	nanaaaacan	aaactgatct	tttgcataac			400

<210> 2

<211> 148

<212> DNA

<213> Mus musculus

<400> 2

cgggtgctag	aagacgacag	aaggggactc	actcgttttg	agaagaccat	ganagcggca	60
gccntctcca	tgccaagggt	gaacaaaatg	ccaggaaatg	tcttttctgc	tancccaaag	120
gattcgaaan	aacacagcca	ttctcttc				148

<210> 3

<211> 400

<212> DNA

<213> Mus musculus

<400> 3

tacggctgctg	agaagacnac	agaaggggaa	ncagctgcgg	tgcccgcggg	agtctgacaa	60
tgcaaatgct	catgttcctg	gctgtccngc	acgactgcgt	acccatggac	aagagtgcac	120
gcaacggccc	caaggtcnan	gagaagcggg	agaaaaatga	tcggacactc	ttnaangatt	180
ggaanacccg	tttgagctac	ttctngcaga	gnncctctgc	tcgccgggag	cccnnaactg	240
gcaataaaaac	naacagcata	cttttatcna	nccttctcct	nangaagcgc	anctctnngc	300
ataaactttt	natnanctgc	tgccngtat	atatgggctg	gctgcattca	nggcgttttt	360
taaagtccca	gttctntgaa	gaaaacattg	aaattctggt			400

<210> 4

<211> 400

<212> DNA

<213> Mus musculus

<400> 4

```

gcctgtggtg gctccgga cctcccttc ctgctcatcc tgggtgtgtac gs ccggag 60
cacctgaacc agagaaacgc cattcgggca tcttgggggt ccatccgcga agcccggggt 120
ttcagagtgc agacgctctt cctcctggga aaacctagaa gacagcagct tgctgacctg 180
tcctcagagt cagcagcaca cagggatatc ttgcaggcct ccttccagga ttctaccgc 240
aacctcacc ccaagaccct cagtggactg aactgggtga acaatactg tcctatggcc 300
cgctacatcc tcaagacgga tgatgacgtg tatgtcaacg tccanancg ggtgtcagag 360
ctgatacaaa gaagggggcc ttcggacaat ggcagaagg 400

```

<210> 5
 <211> 1081
 <212> DNA
 <213> Mus musculus

```

<400> 5
cccacgcgtc cgatctcctc caggggccacc aagcacctct gaagagccat gttccaagct 60
gccggagccg cccaggccac cccctctcat gaagccaaag gcagcagtgg cagcagcacg 120
gtacagcggg ctaagtcctt tagcttgccg gctcagggtga aggagacctg tgcagcctgc 180
cagaagactg tgtaccgat ggagcggctg ttggcagaca agctcatttt ccacaactct 240
tgtttctgtt gcaaacactg ccacaccaa ctcagcctgg gcagttatgc tgcaatgcac 300
ggtgaatttt actgcagacc tcactttcag cagctgttta agagtaaagg caactacgat 360
gaagggtttg gtctgtaaca gcacaaggag ctctgggccc acaaggaggt ggactcaggc 420
accaagacgg cctgagaccc ctttaacacc cattccctcc cagcacatgg cctccgctg 480
ggcagtggaagg aggagattaa cccgggggcg cggggtggga gaggatgagg ctccctcaca 540
caggtttcag gcataaggct ctgctccagg attccttact ttcccatgg gaggttgccg 600
ttgggaacca gaattggaat ttccaccata ctgtgtcctt tagtccacct catctcacc 660
cacggctccc tgggaggccc acaagcccag cttccatact taggtgtctt tctccagcaa 720
ggagtgcaga tgcctcctc aggttcccaa gctccctcac tgcacacctg gccttgtgta 780
ccccctgtgc tccccatcta cctctgcccc ttagcctggg aatgagccac agagactgga 840
agaggggagag tgccatntac tgggcctcat agatgccacc tcgctgaggg gggagggtg 900
gggaagaggg aagacagcct gcagccttca gggctctggg gtcccttgca ccacaaagct 960
aaagctcttg ctagagcctc agctgacagg gtcggcagta gctatgctcc tccatctgtt 1020
gtgctgttct gttgtgatca accctctttt aaaaacattt aaacagcaaa aaaaaaaaaa 1081
a

```

<210> 6
 <211> 743
 <212> DNA
 <213> Mus musculus

```

<400> 6
attggtacgc ctgcacgnac cggtcgggaa ttcccgggtc gaccacgcgt ccgcccacgc 60
gtccgatctc ctccaggggg caccaagcac ctctgaagaa catgttccaa gctgccggan 120
gcgcccangc caccctctct cntgaagccc aangcagcan tggcagcagc acggtncagc 180
ggtctaagtc ctttaacttg cgggctcagg tgaanganac ctgtgcagcc tgcengaaaa 240
ctgtgtncce natggancgg ctggtggcag acaancctnt ttccacaac tctgtttct 300
gttgcaaaaca ctgccacacc aaactcancc tgggcagtta tgctgcnatg caccgtgaat 360
tttactgcag acctcacttt cagcagctgt ttnagaatna aggcactac natgaaagg 420
ttggtcgtta acagcacaan ganctctggg ccncnaggga agtngactca ggccnnaan 480
aggctganaa cctttaaaca cccattccct ccnncacat tgnctccnc tnggcagttg 540
gaaaagaaaa taanccnng gcncgggtt ggganaagaa aaaggtcccc ccncgggtt 600
ttccggcnta agggctctcc ccnnaatccc ttcttttnc cctgggaagt ttgggtttg 660
gaacccaaat tgggaatttc cctnccgtn ttctttntt cccnccnct ccccccnng 720
gtccctggga aggcccaan ccn
743

```

<210> 7
 <211> 743
 <212> DNA
 <213> Mus musculus

```

<400> 7
gggggagttc tgagtcgcat cggccgttnc gtaagcttgg atcctctana acgggcgccc 60
tttttttttt gttttgctgt ttaaatgttt ttaaaaagaa ggttgatcac cacagaacng 120
cacnacngat ggaggagcat anctnctgcc gaccctgtca gctgangctc tagcaanaac 180
tttagctttg tgggtgcaagg gacccnnac cctgaaagct gcaggctgtc ttgctcttc 240

```

```

cccccccctc ccccccctc gaagtggcat ctatgaagcc cagtanatgg ca ctccct 300
cntcctttct ctgtggctca ttaccangct aangggcnga ngttnatngg ganacaaagg 360
ggtnacacnat gcccangtgg cantnaagga ctgtggaccc gaagaaggca tgctgactcc 420
ttgctggana aaaagcacct aanttggaaac tgggcttgtg ggcctcccnc ggaaccntgg 480
ggtgagatna agtggactat aggacctat ggtgaaaatt ccaattctgg ttcccaacct 540
nacctcccat gggaaaaant taggaatcct gcaaccaaac ccttatgcct gaaaaacctgt 600
tttnaaggga ncctcatcct cttcccanc ccgngccccc nggggtttaa agggaaaaag 660
ggtttttttn aaaggggggt ccccnnggcc ccccnntggt tncctnaat ttcccccccc 720
cctnnttttg gcccnaaaaa ccn 743

```

<210> 8
 <211> 174
 <212> DNA
 <213> Mus musculus

```

<400> 8
ggattccggt taggaatgaa attagaagct gtagatctca tggagccacg gttaatatgt 60
gtagccacag ttactcgaat tattcaccat ctcttgagga tacattttga tggttgggaa 120
gaagagtatg accagtgggt agactgtgag tcccctgacc tctatcctgt aggg. 174

```

<210> 9
 <211> 711
 <212> DNA
 <213> Mus musculus

```

<400> 9
ggggctcagg gaaggccgat cttccgggtg gagggggaag cggcgtgact ggagtggaaa 60
tttttcccaa cacaacttct cggaggcaac atattggaag ggactcgggg aggccggagt 120
ccaaatggaa gtggctgaaa gaaacttctc gccctgctga ttctgagccc cgcgtcgtgc 180
cgcgcgccct caattacctc atcgacctgt ggtcttgaca gaacattctt cacaatccaa 240
aagaaaaagc agactggttc gggatgtttg acggctatga tagctgcagt gaggacacaa 300
gtagcagctc cagctctgag gagagtgaag aagaagttgc tcctttacct tccaatctcc 360
caatcatcaa gaataatgga caagtctaca catacccaga tggtaaatct ggcatggcta 420
cctgtgagat gtgtgggatg gtcggtgtgc gagatgcttt ttactctaaa acgaaacgtt 480
tctgcagtgt ttctgttca agaagttact cgtcaaaact taagaaggca agcattctgg 540
ccngacttcn ngttacgggt tngcctccca cnaagaaagc caaantcctt ccnaaacnac 600
cnttagttgg taaattnnct gcctatgccc ntntccanct accttgcnna atccnnccag 660
acnaaancng gcnattctgc catctctgtg gaaagggtcc ctgggggtta c 711

```

<210> 10
 <211> 356
 <212> DNA
 <213> Mus musculus

```

<400> 10
tacggccgag anaagacgac agaagggtnc gggctgcgan aagacgacag aaagggggcc 60
tttctccgct ggccccggcg cgcgccgacg ctectccccg gccatggcgt tcactttcgc 120
ggccttctgc taratgttg cgtgctgct caccgccgag ctcatcttct tcgccatctg 180
gcacatcata ncnittgatg agctgaagac cgactacaag aaacctata gaccagtga 240
ataccctgaa cctcttgctc ttccagaagt actcatccac gcgttcttct gtgtcatgtt 300
tctctgtgcg cagaattggc tgaccctggg cctcaatatg ccccttttgg catacc 356

```

<210> 11
 <211> 400
 <212> DNA
 <213> Mus musculus

```

<400> 11
tacggctgcg agaagacgac agaaggggag tgctcgccgg tcccagcgtt gccataccat 60
cgagatggca tctcgtagcc ggagaccgga acacagcgga ccgcccggag tgttttatga 120
ccagaatgaa gcccggaat acgttcgcaa ctcacggatg attgacatcc agaccaanat 180
gactgagcga gcgctggagc tcctctgttt accaganggt cagccttctt acctgttaga 240
cattggctgc ggttctgggc tgagtggaaa ttatatctca gaagaaggac actactgggt 300
gggcattgac atagccctg ccatgttggg tgccgccttg gacgagatac anaaggggac 360

```


tgctgtctang gacttggg: aggtctcttt caaacccggcc 400

<210> 12

<211> 400

<212> DNA

<213> Mus musculus

<400> 12

acggctgcga gaagacgaca gaagggtacg gctgctgagaa gacgacagaa gggtagcggt 60
gcganaagac gacagaaggg tacggctgctg agaagacgac agaaggggga ggaaactgta 120
aaagctatgc tggggangca gcgcgganct tgattcacct tcacctgctc tggccaccgg 180
ctgaccocgg gtttccggcc gganagcagt canatatgaa cggacgggtg gattatttag 240
tcacggagga agagatcaac ctgacgagag gaccttcggg gctgggcttc aacatcgten 300
gtgggacaga tcaacagtat gtctccaacg acagtggcat ctacntcagc cgtatcaaan 360
aagatggggc tggngccan gatggcggc tccaggaagg 400

<210> 13

<211> 1360

<212> DNA

<213> Mus musculus

<400> 13

agctctagtc cccagagatg tcgccactac tgctgtctgct gctgtgcttg ctgctgggga 60
atttgaggcc tgaggaggcc aaactgatcc gtgtccctct tcaacgaatc caccctggac 120
acagaatctt aaacccactg aatggatggg aacagctggc agagctttct aggacctcca 180
cctctgggtg caacccctcc tttgtgctc tctccaagtt catgaacacc cagtatttag 240
gaactatttg tttgggaacg cctcctcaga atttcaccgt tgtctttgac acgggttctt 300
ccaactttgt ggttccgtcc acgagatgtc atttcttcag tttggcatgc tggtttcacc 360
atcgctttta tcccaaggcc tccagctcct tcaggcccaa tgggaccaag tttgccattc 420
agtatgggac cggcgcgctg agcggaatcc tgagccagga caatctgact atcgggggga 480
tccacgatgc ttttgtgaca tttggagagg ctctgtggga gccacgctg atctttgctt 540
tagcccaact tgatgggac ctgggctcgt gcttccccac tctggctgtg ggcggagtgc 600
agcctccgct ggatgcgatg gtggagcaag ggctgctgga gaaacccgct tctcctttt 660
acctcaacag ggattctgaa ggggtctgat ggggagagct ggtcctaggg ggctcagacc 720
ccgctcacta cgtacctccc ctacacctca taccagtcac catccctgcc tactggcagg 780
tccacatgga gagtgtgaag gtccgcacag ggctgagcct ctgtgcccag ggctgcagt 840
ccatcttaga cacaggcaca tccctcatca caggacctag tgaggagatc cgggcttga 900
ataaaagccat tgggggatat ccttctctga atgggcagta ctccattcag tgttccaaga 960
cgccaacgct tccccctgtc tcttccacc ttgggtggagt ctgggttaac ctccacaggcc 1020
aggactatgt catcaagatt cttcagagcg atgttggcct ctgcctgttg ggcttccaag 1080
ccttggtat ccccaagcct gcgggacccc tctggatcct tggggacgtc tttttggggc 1140
cctatgtggc tgcctttgac cgtggggaca aaaacgtcgg ccgcgcgtg ggactggcgc 1200
gtgctcagtc tgcctcaaca gaccgggacg aaagaaggac tacgcaggcg cagttcttca 1260
aaagacgccc tgggttaggg acaagctcac cgggcccacg cagctatgct tctttccaat 1320
taaacaaact aaaaaaaaa aaaaaaaaa aaaaaaaaa 1360

<210> 14

<211> 1712

<212> DNA

<213> Mus musculus

<400> 14

tcttcgaaa cggggctgag ggggaatctg gacaggggaa tcctggacgt ggagatcgtg 60
agtcattgct gcttcccgag acgctgatga gatccacaag gacgttcaga actactatgg 120
gaatgtactg aagacatctg cagacctcca gactaatgct tgtgtcacgc gagccaagcc 180
gggtccccagc tacatccggg aaagtctgca gaatgtacac gaagacgtta gttcgaggta 240
ttatggctgt ggtctgactg ttcctgagcg gctggaaaac tgccgaattt tggatctggg 300
tagtggggag ggcaggatt gctatgtgct tagccagctg gttggtgaga agggacatgt 360
caccggaata gacatgactg aggtccaggc cgaagtggct aaaacctatc ttgaacacca 420
catggaaaaa tttgggtttc aggcacccaa tgtgactttt ctccacggcc gcatcgagaa 480
gttggcagag gctgggatcc agagtgaag ctatggatt gtcataacca actgtgttat 540
caaccttggt cctgataaac acaagtcct ccaggaggtc tatcgagtgc tgaagcacgg 600
cggggagctc tatctcagtg acgtctatgc cagccttgaa gtgccagaag acatcaagtc 660
gcacaaaagt ttatgggggg aatgcctggg aggcgctctg tactggaagg atcttgccat 720

```

cattgcccaa aagattgc tctgccctcc acgtttggtc actgccgata tc actgt 780
tgaaaacaag gagctcgaag gggttcttgg tgaactgtcg tttgtgtctg ccacatttcg 840
cctcttcaaa ctccctaaga cagagccagc cgaagatgc cgagttgttt acaatggagg 900
aatcaaggga catgaaaagg aactaatttt cgaatgcaat ttcacattca aggaaggcga 960
agctgttgca gtggatgagg agacggcagc tgcctgaag aactcacgtt ttgctccgga 1020
ttttctcttc acacctgttg acgctcgcct gccagctccc caggggcccgt tctgagttag 1080
agacaaagg tctaatacga gatccattca agcttgacga ggactctgac aagatgaagc 1140
ccagacatgc acctgaaggc acgggaggct gctgtggcaa gaggaanaac tgctagatct 1200
acagccagcg cggagcccac cgggtcaag aggttggtta aaggacagtc acagaggctt 1260
cttagcctgc tcttcgccag tgcacagatt atgtgaagg ggcaagcca ccacaagcta 1320
gaccactgct aagaataaga gtgactttta gaggatgtta attgaagggt cacagcaaat 1380
cgctgcttt tctatttctc tatctcagag tctgtgtgcc acctagtggt cagaagtata 1440
acttggaagc ccaaggttta ctcaaaggc caaaggcatc atcaacgttg tgagaattat 1500
cttcctcttg gcctaccaca ggacacctct ggttctctct ctgtgtgtac caggaagcac 1560
agtacttact aaatttatgc taacctgac aaaagattgt caactcaaat ttgtaggag 1620
tattctttag gttgtgtct gcaatttttt tctctgtaac tgaatgaaaa agaaaaacaa 1680
taaaaaataa atttgacttc gaaaaaaaaa aa 1712

```

<210> 15

<211> 1539

<212> DNA

<213> Mus musculus

<400> 15

```

cggacgcgtg ggggagtcgt tggagccccc ggtcaaagac ggcacccctc accagcagca 60
cgtaaaagttt ggcaagaaat gctggcgcaa agtgtgggct ctgctgtatg cgggaggccc 120
atcaggggta gctcggctag aaagctggga cgtgcgtgat ggtggccttg gaccagcagg 180
cgacaggctc acagggccca gccgtcgagg ggaacgccgg gtcatacgtt tggctgactg 240
tgtatctgtc ctgcctgcgg atggcgagag ctgtcccaag gacactgggt ccttcctgat 300
taccaccact gagcgaagcc acctgttggc tgcacagcac cgccagtcct ggggtgaccc 360
catctgtcag ctggccttcc cgggtaccgg agaattgtcg tcaggatcag gacaggctga 420
gaatccaaaa aggggctttg ttcccatgga agaaaactct atctactcct cctggcagga 480
agtgaaccgaa ttcccggtga tcgtgcagaa gacagaagcc acctcccgct gccagctgaa 540
aggaccctac ctctggtgct tgggccaaga tgacatccaa ctgagggaga catcnaagcc 600
ccaggcctgt tttagctggc cctaccgttt cctgcgcgaag tacggctctg acaagggtgt 660
gttctcgttt gaggctggcc gccgctgtga ctcaggtagg ggcctttttg ccttcagtag 720
cccgcgtgcc ccagacatat gtggggttgt ggctgcccgc attgcccgc agcgggagcg 780
tcttccagag ctggccatgt cccaccctg cccctgcct cgggcccctc cctgcctc 840
cctagagccc cctggagagc ttccgggagg gggcccagga tttgagctgc ccactcccag 900
aaagctgcct ctaactgatc ccggccctca aagcctacca ttgctgtca gcccacaac 960
aagaaggacc ggcacccgt ctctatgcgt ccgtgtgcaa gcagaccagc aagcacacag 1020
gcacggcgga gcatttctat gagaacgtgt gcatgctgga ggccagcctt gggctgacca 1080
atgggggttc tgaagcccaa gaggggcccc ttggtggccg cagccccctt gggcagcctt 1140
atctaccata aactgagga tctgagttgg ccgggctcgg ccaggacag caatctggaa 1200
gcccagtagc ggaggctgct ggaactggag ctggatgagg ccggaagcgc cggccgctct 1260
ggagcgcagg caggcatcaa ggccaagctg gtgacctgc tgacctgtga acggaagaag 1320
ggccccgccc cctgtgaccg gccctgaagg cctgagcggc cagccactgc aggacagag 1380
tgatcaccca agaccaggaa caacttcgaa cataaccctg ctactctgac ctgcagggac 1440
aagccagggtg gcccggggag gagccacact ctgccctacc tccctccctc gactgtacag 1500
attgaacagt aataaagctt gcctatcaac ttcaaaaa 1539

```

<210> 16

<211> 3599

<212> DNA

<213> Mus abbotti

<400> 16

```

agagacagcg tgatcccgcc ctcccacggg gcagctttta ctgtctaggg aagaaatccc 60
caaagtccat ggagctctgaa gactctgtca agcctcgtca ggaacactag gacttttaga 120
gggcacttgg caccggaagc tagccgggta ggcggagcct cacttgatt gacttcacag 180
ctgcctagac aggcctcagc taggtgctgg gcacctggga ggaggaggag acattagcag 240
caaaggctgt taacagaagt gcctgcctag gcttgaggc aagacgctgc tgttcacagt 300
gcgagacgga ggtaggagta taatggctgt ccagggtctg cggcagatgg tctacttctc 360
actgagctctg ttttctctgg tgcaagggtg acacagtggc agcccccgag aagacttccg 420

```

```

cttctgtggc cagcggaa agacccaaca gagcaccctc cactatgata aa ttcaga 480
gcctcacatc ttgtgtgga acacagagga gaccctcaca attcgtgccc ccttcctggc 540
agccccagat attccccgct tcttcccaga gcctagaggg ctctatcact tctgcctcta 600
ctggagtgcg cacactggga gactccactt gcgctatggc aagcatgaat acctgcttag 660
tagccaagcc tccagactcc tctgcttcca gaaacaggag cagagcctga agcagggagc 720
cccgtgatac gccacctctg tcagctcctg gcagattccc cagaacacca gcctgcctgg 780
ggctccgagc ttcatcttct ccttccacaa tgccccacac aaggctctcc acaatgcatc 840
tgtggacatg tgtgatctca agaaggaatt gcagcagctt agcaggtacc tgcagcacc 900
tcaaaaggct gccaaagcgc ccaccgcagc gttcatcagc cagcagttac agagcctgga 960
gtcaaaagctg acctctgtga gcttccctgg agacacatta tcttttgagg aggaccgggt 1020
caatgctaca gtgtggaagc tgccaccac agccggtcta gaggatctgc atatccactc 1080
ccagaaggag gagggagcaga gtgaggtcca ggcatactcg ctgttgcttc cccgggccc 1140
attccagcag accagaggcc gtcgcccggg tgacgccaag aggtcctctg tagtagactt 1200
cagcagccaa gctttgttcc aggacaagaa ttctagccaa gtccctgggtg agaaggtctt 1260
gggtattgtc gtgcagaaca ccaaaagtcac caacctctca gatccggtgg tactcacctt 1320
ccagcaccag cctcagccaa aaaatgtgac tctgcagtgc gtgttctggg ttgaagacc 1380
ggcatcaagc agcacaggga gctggagcag tgcaggctgc gagacagtga gcagagacac 1440
acagacatcc tgctgtgca accacctgac ctactttgca gtgtgatgg tgcacccac 1500
agaggtagaa gccactcaca aacactacct cagctcctg tctacgtgg gctgtgtcat 1560
ctctgctctg gcttgtgtct tcaactatcg tgccctacct tgcctcagga ggaagtcacg 1620
tgactacacc atcaaaagtc acatgaacct gctgtccgct gtcttcctgc tggacgtgag 1680
cttctgtctc agcgagcctg tggcactgac gggctccgaa gcagcctgct gcaccagtgc 1740
catgttctct cacttctccc tgcttgctg ccttctcctg atgggctcag agggctacaa 1800
tctctaccga ctggtgtggg aggtcttcgg tactatgtg cccggctatc tgcacagct 1860
gagcatcgtg ggctgggggt ttctgtctt cctggctact ctgggtggcg ttgtggatgt 1920
gaataactac ggcaccatta tctagctgt gcgcccgaat ccggaacgtg tcacctaccc 1980
ctctatgtgc tggatccggg actccctggg gagctatgtc accaacctgg gcctcttcag 2040
tctggtgttc ctgttcaacc tggctatgct ggccaccatg gtggtgcaga tctgcggct 2100
tcgcccgcac agccagaact ggccccacgt gctgacctg ctgggctcga gcctgggtct 2160
tggtctcccc tgggcttgg tcttcttttc ctttgcttcc ggcaacctcc agcttgtcat 2220
cctctacctc ttccagatca taacttctta ccaaggcttc ctcatcttcc tgtggtactg 2280
gtccatgcgg ttccaggccc aaggcggccc ctccctctg aagaacaact cagacagcgc 2340
caaaactccc atcagctccg gcagcacctc ctccagccgc atctaagcca ccgccacacc 2400
tccccctccg gaggacacat gcatggcgtc cgctcacgat gtctgtggcc cagtgtctgt 2460
cccacccagc cttgtgttgt tagtggcata ctagaagaag ccctggtcct tgaaggcgta 2520
gggctgttgt tctggtaggt agatacctag cttgccttgg ggacgactct ggtcctcaaa 2580
gggctcagaa gcacactgcc attctgtttg tggggccgtt tcagtctgga gctaaggcct 2640
tgtctttctg gccacctctg ggtccagctg ttgctgctgg gtgttgagac ctgcagaccc 2700
aagctggggg tagatctcga aggaggctga cacatccggc ctgagacaca gctaactgtc 2760
ttgacttgct gctctgtctc tgtgttcacc atgcagatcc cgagggtggc actgggggta 2820
aatgttctgg gagaaggttt ggaggcagag cacttagga gctgagcatc tccccagcc 2880
tttctgcaaa cctcctctct cattccccat ccccaacccc tctctgtgt tccccaaac 2940
ctccacctga agcctggggg cctagaccaa tgcgtgtgatt tggggtggtta gttccagca 3000
gtttcctggt gccagctatc aaacttctgtc tgttgtgtgg gctttggcct ctgactcagg 3060
gcagggttct gctgagccc tctctccaag ctgcctcacc ttgtctcgca cctcagaggg 3120
acctccatct cctctgaagc ctctcctctc tggcaagtac tgggatacag ccacctcttc 3180
aaccagcac tctgaagacc aagacagccc cctctggtga cactggccaa gcttgatctt 3240
tttctaaga agtggctctc agatccccgc aggtcgctca gaagacactg ggctgcctag 3300
tgtgaattct gtcctactaa cgtacagtga gcagctctc acccccaccc ccgcaaaagc 3360
tctaccaag tcttgagtg tcaggcaggg ggctggaaat ccaggaggac ttctgcaaa 3420
aggcagcatt tcatcttgac ctcagccttc aggttgggga gaatgttct tttaaatacc 3480
agttcatttg tcttttgata ttaaagctct ttatagagag tctggaaact gtaggcgatt 3540
gtcagaaga gaaataaaaa tgagctgtta tctaagcca tggcaagca gcacaaaaa 3599

```

<210> 17
<211> 399
<212> DNA
<213> Mus musculus

<400> 17
agcgacatgg ccccgcccg gctccaggcc cagcctccag cgggctctca actgaggttc 60
ctgctgttcc tgcctgtgtt gctgctgtct ctgtcatggc catcgaggg ggacgcccctg 120
gcaatgctcg aacagcgacc ctccggccct gagtcccaac tcaacgcccga cgagctacgg 180
ggctgcttcc aggacctgct gagccggctg catgccann atagcgagag gactctaact 240

```

ctgaaccaag tcccgac: ctgtcnggat actcantcca cangtgagat tc jtccac 300
ggccagctgc taccnccgt caaccggcg tcgctganc anggtctccc cgaacctacc 360
ncgtgcancn agcgctgctc ctgctgacnn cnaangccg 399

```

```

<210> 18
<211> 400
<212> DNA
<213> Mus musculus

```

```

<400> 18
cggacgcgtg gggaaagcgg gtggggatat acgtgtgctg gtgccctaca attcgacagg 60
cgtcttgga gggcgacca ccttgcactg tactctgact tctaatgaga atgtgactat 120
cactcaaaata acctggatga agaangattc aggtggaccc cacgctcttg tggctgtctt 180
ccaccccaag aaggggccca acatcaaaga gccagagagg gtgaaattct tgggtgcccc 240
acaggatctg aggaacgcat ctctggccat ctggaactta agtgttnnaag acgaangcat 300
ctatgaatgt cagattgcca cattccccag aagcagtana agcaccaatg cctggctgaa 360
ngtgcaagcc cgacctnana aactgcana agccctggac 400

```

```

<210> 19
<211> 2313
<212> DNA
<213> Mus musculus

```

```

<400> 19
cgaattcctg ggtcgacca cgcgcccgag agggtagaca gaaaggcggg aaagggtctgt 60
gaggtcaaat ggaccccatg gaactgaaca acgtcagcat cgaacccgac ggagacagct 120
gcagcgggga cagtattcag gacagctaca ccggcatgga aaactccgac aaggacgcca 180
tgaacagcca atttgcta atgaatgccc aaagtcagaa gttcctgaca aatgggtttt 240
tagggaaaga gaagctagcc gattacgcgg atgagcatca ccctggaatg acttcctttg 300
gaatgtcctc atttaacctg agcaacgcca tcatgggcag tgggatctta ggctgtctct 360
atgccatggc caacaccggg atcatccttt ttataatcat gctgcttact gtggcaatac 420
tctcgctcta ctcggttcac cttttgctga agacagccaa ggaaggaggg tctctaactc 480
atgaaaaatt gggcgagaaa gcatttggat ggccctggga aattggagcc ttcattctta 540
ttacaatgca gaacattgga gccatgtcaa gctacctctt catcattaag tacgaactgc 600
ctgaagtaat cagagcattc atgggacttg aagaaaacac tggggaatgg tacctcaacg 660
gcaactacct cgtcttattt tggtccgtgg ggatcatcct cccgctctct ctccttaaaa 720
atttaggcta ccttggttac accagtggat tttctctctc ctgcatggtg tttttcgtca 780
gtgtggtgat ttacaaaaaa ttccaaattc cctgcctctt gctgctctg gatcacaaac 840
acggaaatct gacgttcaac aacacacttc cgattcacat gatctcgctg cctaataact 900
cggagagctc ggggtgtaac ttcatgatgg attacgctca ccacaacca gctgggctg 960
atgagaagca ggtcgagggc cctcttcaca gcaatggcgt ggagtagcaa gccacgggtg 1020
ctgagaaatg ccaacaaaaa tactttgtgt tcaattcccg gacggcctat gcaatcccaa 1080
tcctggcttt tgcttttgtc tgccaccctg aggtccttcc catctacagc gagcttaaa 1140
atcgatcccc cagaagatg cagacggtgt ccaacatttc catctcaggc atgctcgtca 1200
tgtaccttct tgcggccctc tttggttacc tgagcttcta cggggacggt gaagacgagc 1260
tgctgcatgc ttacagcaag gtctacacat ttgatacggc tcttctcatg gtgcgcctgg 1320
cagtctggtt ggcagtgaca ctgaccgtgc ccatcgtgct gttcccgatc cgtacttcgg 1380
tgatcacact gctgtttcca aggaaccctc tcagctggct gaagcatttc gggatcgctg 1440
caatcatcat cgcactcaac aacatcctgg tcatcctcgt gctaccatc aaatacatct 1500
ttggattcat aggggcttct tctgccacta tgctgatttt cattcttccg gctgcgtttt 1560
atctcaagct cgtcaagaaa gaacctctaa gatcacccca gaagattggg gctttggtct 1620
tccttgtagc tggaaattatt ttcatgatgg gaagcatggc gctcattata ctgactgga 1680
tctacaaccc gccgaatccc aatcccgagg agacgcgtct cactggaata 1740
cagctgaaat tgcctgaagg acattttagt tgtcttgatt gggatgttag tctgaggaat 1800
tagcaagatt ccaaagacgt ttttctagct ctatcagcac acattttaac ccaggccctg 1860
cagtgcagtg tctgatgccc gagttgtgtt tgcagcagct gtgcaagctg aagcctgttg 1920
gctgcgtgtg ttggtcagca gacaatagcc tgtcccccca tggtaactcc acttctctcc 1980
acccccagat taacaggtaa ttctactctc agaacatcag acaaagacct cctggttggg 2040
atacttgtag aagagaaaaa tatgggtttt gttgggaatg gttttgttgg gaatgggtgaa 2100
ggatgcatta aaaattctgt gcgaagtac atcagttacg gccatctctc actctacacc 2160
aacactaagg gtcgggtgac tagctgaggg agggggatat cttgggctgt ccctgtgagg 2220
atcatgacgt atgacggttg ccagtataga gtacttcatt tcaatactca aggaatagtt 2280
tgcccaacct gcttattaca ccgagttagt gaa 2313

```

<210> 20
 <211> 3408
 <212> DNA
 <213> Mus musculus

<400> 20
 gcagctaccg gtcgggaatt cccgggtcga cccacgcgtc ccacggctgc gagaagatca 60
 cagaaggggg cggtagacga ggaccaggac tcgcgcgtcc agcggagaag caggagaagc 120
 cggcgacctt gcgtctctag cctgatccct gtcttgccgg cctgaacatt cgcagctgga 180
 gagatggcgt tcgtgaagag tggatgggta ctccggcaga gcaccattct gaaacgctgg 240
 aagaagaatt ggttcgacct gtggtcagac ggtcacctga tctactacga tgatcagact 300
 cggcagagca tagaggataa ggtccacatg cccgtggact gcatcaatat ccgcacgggg 360
 catgagtgcc gggacatcca gcctccagat gggaaagccca gagactgtct gctgcagatc 420
 gtttgccgag acgggaagac catcagttct tgtgcagaga gcacagacga ttgcctggca 480
 tggaaagtta cactgcagga ttccagaaca aacacagctt acgttggttc agcaatcctg 540
 tctgaagaga ctgcagtggc cgcgtccccc cctccctacg caacctatgc tacaccnacc 600
 cctgaggtct acggctatgg tccatacagc ggcgcatacc ccgcaggaaac tcaagttgtc 660
 tatgccgcca acgggcaggc atatgcagtg ccataccagt acccgtatgc aggagtttat 720
 ggacaacagc ctgccaacca agtcatcatc cgcgagcggg accgagacaa tgacagtgc 780
 ctggctctgg gcatgctcgc cggggcagcc accggcatgg ccttgggctc tctgttctgg 840
 gtcttctaga gccttcaaca ttttctgtgc atagcttctg ttagtctctg gtgcagtaat 900
 ttgatttgca gggcatttct gtttctgaca agtgtcttcc ataataattt aaatagtctc 960
 tttgaagggt gtaattctaat aattgtgact gacctgcagt gtaccacaaa gaaagcccg 1020
 ggtatgctgt gagtgcagc ctgagtcctt cggggtacta gcttgcacca agtctttctt 1080
 agggactttt ggtatggctt atgtaaacac acccagttaa atgggcaatt tccgtccagt 1140
 taggtgcagt gttgaattaa gggatggctt tcttgctat gccaaacta atactgctga 1200
 tggaggaaga tgtgtgcaag tgtggtgagg agagtacacg ctctcttaac tgtggttct 1260
 ctcttagacc cctgctcgtt gttaccctag gactgtggg ctggtggctc ctgcaagact 1320
 atgggtgtgag gacctgttaa cgtacctctt ggagcactta ggtaccctt gaagctccta 1380
 ggtatcacca gcaggattgg ctgctcagga tgcagagggc caccctccct ctttaaaaat 1440
 tacgctccag taatctgccc agttttattt tcttgttatt ctctgtttg ctttctctg 1500
 ggtatgattg cattagtctg gagttaggaa ttgattcagag tgccggtggg tggaggcagt 1560
 caggagctg tccagcgacc tgcctcagat gtttgtttta ggtatattga ttgccagctc 1620
 aggcctgcaga gagcctatag agactatttt tctacttgta aagaaagtat ggtgagggga 1680
 attgaggaga gccrtgtttg aatgttctct cctcaggcct cctggggccc acactgcgtg 1740
 gtccctggga ggccttttct cagcagtggg agggaggccc cgtggctgtc cagagtctca 1800
 ggttttgagt gagaatggg attgggtaga gcatctcagg gattgtttct aatccctcat 1860
 gttatgggga tccagccgtg ttctcagtc agaccgcgtc acctcagaag agcttaaaac 1920
 atttctggtc cccaaatgtg tggcactctg agaagctcac aatctggctt tctaacgaaa 1980
 atttgtattt ctaaaattag agaatacatg ttccacgcat ttaaaattta tgttctttca 2040
 tgttttaaa ctcccaatc cagctttgtg actggcatat tttagtttca aacagtacc 2100
 cggcacaaag gtgggatggc acagtgaagg cccccgcctc tctactttgc atagtctgt 2160
 tctccagggt tgcctccagg aagcattcat tctgactttg ctccagccag tgcagcgtg 2220
 ctgcttgcc gccgtgctgc tgggtagctc tttcttggtc agatcaagtc tccaacagat 2280
 ctccatgtga gacagttgcc aagttagatg ggtggtgccc atagtgttt ctgcatactc 2340
 cttggggacc tgttgacacc tgccatttc cagctgacat ttgttttct gtcactctg 2400
 atagatggga tatgtgacaa catggtacgg acgcccgttca gtgtcgttt aataagcatg 2460
 atgctgattt tacatcctgt gctgtatgac tgccatttgc tcacagtgc accattgcta 2520
 aagctccgtg ctctacttac aaaacactaa aaccagtggg tagtgtttca cagtgtttt 2580
 aatttttagag ttagttactg gcattcctaa agccatagag tactgagtca catccctgaa 2640
 gtacttttga aacagaattg tctcctactg tccatgggt gtgccctgcc tgtctctg 2700
 ccccaatggg gctagctgta ccaggcagcc atagttgagc ctgatcatc ctgtcaccag 2760
 tttgacttga ttatataccc agaattggaat acattcttgg gcatctcagt tccctagccc 2820
 tgatcctcat agacgccacc ctttcgatgg cttttgccc gtccacttga cctcagtgc 2880
 tccctcgatt ctgagttag aggggacgac ttgtccagca ttgaggaaca tgtctcctc 2940
 actgagactt aaatgatgat gcagggtcgg aagaggctgg ctgctgacac tgcactgtg 3000
 ctggtgtcat tgcctccta gttctttgat ttaagaacct ttcatatgga agcctgagg 3060
 ctccctcaga tcgtcccttg ccaagaaggc ctggcttagg tcattagtgc ccacagtgc 3120
 cttctggagt gtagcaagtt cctgcgtttg agacagaatg gttcagattt attttctaca 3180
 tctgtgtgtg accccatgca cctctcatt ttgccttcca gtctacgtag atgaagatg 3240
 aaaggcagag gatgcagaca gtcttctttg tgattgtctc tgtattctg ttgcatctac 3300
 cgagcccggt ttccctctgt ctgtgcatac agtatgttta taagtgaact tgttaaaata 3360
 ttaaatgatc actaaccccta aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3408

<210> 21
 <211> 878
 <212> DNA
 <213> Mus musculus

<400> 21
 aaatctcagc tgtgcagcct tctagaacct catgtctaca aagtgcacaa ctttctctcag 60
 gcacaatcca tcatacagtg gatcaatcag tcagagtcag aggaagaact agtcaaaatt 120
 acctcatttt ctgaattcat taacacctta aagaaaaccc acaaatacct aatggaagag 180
 agtttcaaaa ctgagccccc agaaagagtg gaagaagcaa agagaatggc tacatatgaa 240
 gtcaccacag ctctcagctc cttcttgaag tacctcagag aaacacagca gccagacatg 300
 cagctgtgtg tactctccat tgctactggg gtaggctatc agttggtaaa cagtatcttt 360
 cagcatcttc tgggggtgtg tgagttaaac ttctctcttg atcaaacgga aaataacgaa 420
 cataaatacc aagaactgaa aaatatttgc aattacagag cccaggcatt cttgggtgctc 480
 acagccctaa gagccacagt tgaaatcaca gatgtttcta cagaagagaa aggacaacgt 540
 ttgacattaa tacaacaaca tatgggtgca ctgttgtctg aagaagttgc acatgtcctc 600
 aaaaacatg gagaacatca tgactgggaa aggctggaga atgatttgag attactcatt 660
 gagggggact ataaagccac caccattcc ttacaaatgg atgaagttaa aaaacaattg 720
 caaagttnnt gccatgaaaa gaaacagact tataaacaac aaggtaatga aaacagaaca 780
 aaagaaatga tagaaatgg acatttcctg gacttactcc aacgtttagg cctagacaat 840
 tactatccaa aaaaaaaaaa aaaaaaaaaa aaaaaaaa 878

<210> 22
 <211> 731
 <212> DNA
 <213> Mus musculus

<400> 22
 actagactct ggtagagaaa aaagtgcctc ggagtggctg ctgaacaacc tcatgactca 60
 ccagagtgtg gagcttttca naaaaactca cttntaaacg tcaggaaaca gctcattcgt 120
 tcntctaact tatttgtgat gcaagtggaa atagatgtnt atacagctct taaaaagtgg 180
 atgttccttc agctgggtgc ntcttgaat ggggtctttg aaaacagctt ttgacaaaaa 240
 cagatgtctg gttttccaag tggaaagaag actttgaaag gacgacttcc cntgaaactg 300
 aacagggaaa accagntcgc gcccntgtnc cgacatttaa ggctacagta cattatcagt 360
 gatctggcct ctgcaaggat cattgagcag gattctctgg taccttcaga atggctggcg 420
 gcagtgtata aacagcagtg gctggctatg ctacgggctg aacaagacag tgaagtggcg 480
 cctcaagaaa tcaataaaga agancctgag ggaacacagca tgaggtgtgg tcgaaagctt 540
 gccaaagatg gtgngtactg ctggcgctgg acacgcttca atttcgggct tgacctcctt 600
 gtgncttaca ccaatcgata catcatttcc caacgccata cncctgaacca gccatgttat 660
 tggatctgtc agcttacacc tcgaaggagc atagcattta aatgcctctg gcttctttga 720
 cagtagtggg a 731

<210> 23
 <211> 400
 <212> DNA
 <213> Mus musculus

<400> 23
 gacatttaag gctacagtac attatcagtg atctggcttc tgcaaggatc attgagcagg 60
 attctctggg accttcagaa tggctggcgg cagtgtataa acagcagtggt ctggctatgc 120
 tacgggctga acaagacagt gaagtggggc ctcaagaaat caataaagaa gaacttgagg 180
 gaaacagcat gaggtgtggg cgaagccttg ccaaagatgg tgagtactgc tggcgctgga 240
 cacgcttcaa tttcggtctt gacctccttg tgacttacac caatcgatac atcattttcc 300
 aacgccatac nctgaaccag ccatgttatt ggatctgtca gcttacacct cgaagganca 360
 tancatttaa atgccttgg cttctttgac agtagtggga 400

<210> 24
 <211> 2649
 <212> DNA
 <213> Mus musculus

<400> 24
 tgagcgcaac gcaattaatg tgagtttagct cattcattag gcacccacagg ctttacactt 60
 tatgcttccg gctcgtatgt tgtgtggaat tgtgagcgga taataatttc acacagga 120

```

cagctatgac catgatt. , ccaagctcta atacgactca ctataggga a5 .ggtagc 180
cctgcaggta ccgggtccgga attcccgggt cgacccacgc gtccgacggc tgcgagaaga 240
cgacagaagg gggcaccaca gaggaccocct tcaacctaaag gaaacaccca ggcttcgata 300
ggaccatgct gcagaggttg cagaaaaagg agatcagcaa ctttgaatac ctcatgtacc 360
tcaacacact ggccggaagg acctacaatg actacatgca gtatcccggtg tttccctggg 420
tcctcgctga ctacacctca gagatgttga acttgacgaa tcccaagact ttccgggatc 480
tttctaagcc aatgggggct cagaccaagg aaaggaagt gaagtttacc cagaggttta 540
aagatgttga aaagattgaa ggagacatga ccgtgcagtg ccactactac acccactatt 600
cctcagccat cattgtcgct tcctacttgg tccgaatgcc accattcacg caggccttct 660
gctccttaca gggcgggaagc tttgatgtgg ctgatagaat gttccacagt gtaaagagca 720
cgtgggagtc tgcctccaaa gagaacatga gcgatgtcag ggagctgaca cctgaattct 780
tctacctgcc cgagttttta accaactgta atgcagtggg gtttggctgc atgcaggatg 840
gaacgacact gggggatgtg cagcttcctc cctgggctga cggggatccg aggaaattca 900
tcagcttgca cagacaggct ctgaaaagt acttcgtcag cagcaacctc caccactgga 960
tagacctaat ttttgggtat aagcagcagg ggccggctgc tgtagaggca gtgaacactt 1020
tccacccta cttctacggt gatagaatag acctgggag catcactgac ccgctgatca 1080
agagcaccat cctgggcttc atcagcaact ttggacaggt gcccagcag atcttacta 1140
aaccocaccc atccagaaac accacaggga aaaacccagg gcctggaaag gatgcttcca 1200
ccctctagg cctccagagc cactcacagt cctcctcca cagcctgcca gcactgagac 1260
cctctcaggt cacagtcaaa gatagtacc tttctctct agggtcggaa tccccaaaag 1320
gggcatcgg ccacatcgct cctactgaga agtcaatcct ggcagtggag aagaacaagc 1380
tgctgatgcc cctctctggt aacaggacct tcagctgggg ctttgatgac ttcagttgct 1440
gcctggggag ctacggctct gacaagatcc tgatgacct tgagaacctg gctgctggg 1500
gtccctgtct gtgcgctgta tgcccttccc ccacgatgat cgtcacatcc ggggccagcg 1560
cagtggtgtg catctgggag ctgagcctgg tcaaaagtcg cccgagaggt ctgaaactcc 1620
gacaggcctt gtatggacac actcaggcgg tcacatgtct gacagcctct gtcacctca 1680
gcctcctggt gagcggatcc caggatcgca cttgtatcct gtgggacctg gaccacctct 1740
ctcgtgtggc ctgcctgcct gtccaccggg aaggcatctc agccattgcc atcagtgatg 1800
tctcgggaac cattgtcttc tgtgccggag ccacttgtc cctgtggaat gtcaatggag 1860
agccctggc cagtattacc acagcctggg gcccagaagg aaccataacg tgctgctgca 1920
tagtagaggg gccagcgtgg gatgcaagcc acgtgatcat cacggggagt aaggacggaa 1980
tggttcggat ttggaagaca gaggacgtga agatgcctgt tcccaggcag gcagtgatgg 2040
aggagccctc cacggagccc ctaagcccca gaggtcacaa gtgggccaa aatcttgccc 2100
tgagccgaga gctggatgtc agtgttctc tgagtggcaa gccagcaag gcaagtcctg 2160
ctgtgacagc tctggccatc actaggaacc agagcaagct cttggttggc gatgagaagg 2220
gcgaatcttc tgctggtctg ctgatgggta ggagacaagg agctggaggg accgactgta 2280
aagcctggag ccctggggtc ggagcaaca ggctacaggc acaagatgat gtgtagcctg 2340
ggctgcttaa ccagagcaag ttttggggg gctccactcc acacagttct caaggagtcc 2400
ctgatgggtt gcaccgtgta ccctaaacat gtctgtagtc tatgggactt ctgtaagaag 2460
gatcttggta gacactgatg ctggaaaactg acgctgcatg ggaaatttct acgtggctca 2520
cttcaccaag gcttattgca ctgggaaaag aagagggggg gggatttggt tcatggaaac 2580
caccctactg tctttatttt attaaaaactc cattttccga aaaaaaaaaa aaaaaaaaaa 2640
aaaaaaaaa
2649

```

<210> 25
 <211> 499
 <212> DNA
 <213> Mus musculus

```

<400> 25
ccattactt ccacotccca tgaggactgc acctgcggca gcgattttat gtgaacttgg 60
attactgtat ggaaatggag gagtgaagaa gtgtgggaat cataaaagaa ggcagaacca 120
aaagacagag ggatcttggg ccattgagtca cagcaaggaa agcctccac caaacatctg 180
cactggactc ctacaggaac aagaaatgaa ctatcgtgtt cagccctga gacttgggct 240
taacatgatt ttcacattgt ctaccctacc taatagagca gagatgtaaa tattattctt 300
atttttagagg tctgatgcct cagctgcaat gggtgagaac tactcctcat ttattctctt 360
ccaaggcaat aaaagagaat ggacaaaaga cagtctgtca tcacatctag tcaaaagagc 420
taatgtcgca gtacaactct tcaaaagaaa aagaaaaaac aagaaaaaag taataaacag 480
atgtgttctg cttgaaaaa
499

```

<210> 26
 <211> 686
 <212> DNA
 <213> Mus musculus

<400> 26
 cgggggtcgcg cgggaattaa gtgagtcocat tcttgccacac cgagaaaata tgcgacagat 60
 gataagaagt tcttctgaac cctttgggaa gagacttgct cagtatctct gatggtagan 120
 ggagagctca taatcgtaga ggacataatg atgggtgaaga ttctttgact catacagatg 180
 tcagctcttt ccagacaatg gaccaaagtg tgtctaatat gagaaactat atgcagaaat 240
 tagaaagaaa cttcgggtcaa ctttcagtgg atccaaatgg acattcattt tgttcttcct 300
 cagttatgac ttattccaaa ataggagatg aaccgcaaaa ggtttttcat gcctcaactc 360
 aaactcgtcg agtccagga agaataaagg aaaccaggaa agcnatgaga gattctgaca 420
 gtgggactag aaaaaatggc tattggtcct catatccatg aaccgagctc ntgttcntta 480
 ataaagtccn nngancnag aaagactggg agatgaaaaa aggtccaccc aggaanttca 540
 tcccttatga aatgaaaatt gatgccccct gccttttgaa tgaagggaatt ggccaaattn 600
 aaggtttttg aaattttcaa acccgggaac aaccaatcc tnggaaaacc cctnnaaatg 660
 aaaaaatttt ttgggcccct aaaaaa 686

<210> 27
 <211> 651
 <212> DNA
 <213> Mus musculus

<400> 27
 caaaaannacn tgcgttttagc aaataaaaaac tgcccagtg gacaanctgc gccgggaatc 60
 gctgtcagta ctgcccantt canaaatgcc tggctgttgg ganggncaaa gaaatgggnc 120
 gcacagacag nnnaaaaggg cggagaagnc gnntgcccnc gaaaccgaag aaccacanga 180
 nccctctccc ccttcgcccc cgttccangc gaacctgacn ancaaaangaa nggagangan 240
 acccagcana nccngcaant ctatganctc ctgactggct ccanggaaan catcccgggc 300
 ngggcagaaa aaatccctgg gcnnccgaga actgcccacaa gccgaccaag aactgcnnnn 360
 tgaancagcn nctnagaac tgnntgtccn nccannagca nacagaagat gaacatcgac 420
 atntctgcct tctcctgcat ngcngccctg ggctatggnc acnnaaaaac acggcncngg 480
 gaaccggaag anggaanaan gncnccnga tngntgantg ncgcgagac acngacttt 540
 cnacnanggg gggggnannc gcccccttn nttgttccna ctgggggggg gaacnnccn 600
 aaattcngc cctnnnggc ccnagggggg ggggggncn ttttctttct n 651

<210> 28
 <211> 683
 <212> DNA
 <213> Mus musculus

<400> 28
 ggggtctgttt ggactgcaga agcaaggggg tgatgtagcc catccttccc tttggagatg 60
 ctgagggtgt ttcttctgc acccacagcc agggggatgc cactcctccc tccggcttga 120
 cctgtttctc tgcgctacc tccctccccg tctcattccg ttgtctgtgg atggctattg 180
 cagtttaaga gcagaacaga tcttttactt tggccgcttg aaaagctagt gtacctctc 240
 tcagtgtttt ggactccatc tctctctc cantaccttg cttcttactg ataattttgc 300
 tgggaattcct aaattttcaa tgacattttt tttaactact atattgattg tcttttaaaa 360
 aagaaaagtg catattttat caaaatgtgt atttcttata cgcttttctt tgttatacca 420
 ttctctcagc ttatctcttt tatatttgta gganaaaactc ccattgttat gaatccact 480
 gtatgattta taaacagaca atatgtngag tgccttttgc aanaaaaagg gtgtgtttga 540
 aatcatccgg atccagccan gaaactgtct ccaaaggaaa cncnaccctc tctgttctt 600
 gctgttntgc tgatcatccg ccanaaggtg cttcccccg aattttggtt tggttntgtt 660
 tncgaaaaa ttttttccgn ttn 683

<210> 29
 <211> 624
 <212> DNA
 <213> Mus musculus

<400> 29
 aatgcacctc ctgtattcc cactttcgtg ntcatttcgg ttctgatctt gtcaaaccac 60
 gcttgaccgc ttctgacgcc gggatggcct cgttactana cttttctttt taagggaantg 120
 ctgttttttt ttganggttt tcaaaaacatt ttgaaaagca ttactttttt tgaccacgan 180
 ccattgantt tcaaaaaaat ccgggggtgt gtgggttttt ggtttttgtt ttaatttttg 240
 gtgtcgttgc cttttttttt taatgggtgt ggccccatga aatgggtgcc ccactcactt 300
 cctganatc gaacngactg tgaatccgct cttgtcnga anctgaacaa nctgtggctt 360


```

ttttccaact cgggtgtg . tttctgaatg ttatgtggta agaccccgcg gs .ggcan 420
caactgcccc ggaaccccan cccctgcttc catctgtngc tngtgcgccc cacantaaaa 480
cmttcanaen cccctgaaaa gttcttgaaa aaanttaatt anattgtccc nnttttactg 540
ggaaaaaatt anccntactc cnnccaattt tnttcttnen antgggctga aagaaatcct 600
tggtccccan ttttcccttg ggcen 624

```

<210> 30
 <211> 641
 <212> DNA
 <213> Mus abbotti

```

<400> 30
atcccagacc aggcaacaag gtcctttgag gtctataatg aaagatctgc attctgatga 60
caatgaggag gaatcagatg aagtggagga taacgacaat gactctgaaa tggagaggcc 120
tgtnaataga ggangcagcc gaagtcgcag agttagctta agtgaaggca gcgatagtga 180
aagcagttct gcttcttcac ccctacatca cgaacctcca ccacccttac taaaaaccaa 240
caacaaccag attcttgaaag tgaaaaagtc nataaaagcaa agcaaatcag ataagcaaat 300
aaagaatggt gaatgtgaca aggcatacct agatgaactg gtnagagcttc ccagaangtt 360
nargacattg agagaaaagac acattctgca gcagatcgtg aaccttatng aagaaactgg 420
gacactttca tatcaccaac acaacatttg attttgatct ttgctcgtcg gacaaaacca 480
cagtcctgta actacagagt tcctggaaac atctggaaca tcctgaagat ataaccactg 540
gatgcctcna gaactaatgt gtttnnnnnn nnnnnggggt nnnnnntttg ggtgtgaatt 600
ttgttccccg ttgtttaaat gaaaaccccc cnaatgatgc n 641

```

<210> 31
 <211> 400
 <212> DNA
 <213> Mus musculus

```

<400> 31
gcgagccgac ggagnggctc tacggtggat agcgtgttcc ggaacctatc cctcgaatta 60
gccgagtcag gcagagaggg ggcgggngt gcttccgccc ttgctaggag gggctgcatt 120
gcaggggaga cccagcggca gattctgtca cagacgaggg agaaggcgtg aggagacaaa 180
gccgtcacat ccgcgacagc ttcttccagc agcctcctcc tctccagtcc agagccgacc 240
cccagacccc tgaggcatcc ctcccgctct cggaaacaccc tagtcattca ttgctaacag 300
gggatcatga gggacctgtg aagtagccag tacagctcct ttcttttctg gangatgccc 360
atcccacaac tggatctgtc ngagctggaa ggccctgggccc 400

```

<210> 32
 <211> 400
 <212> DNA
 <213> Mus musculus

```

<400> 32
acggctgcga gaagacgaca gaagggaaag taagggctgg tcgctgcgtc tgaacgccgg 60
gttggaagaa ccgctgtgag cccttggcgg acgtnagcgg aaagaagatg gcggtgcagg 120
tggtgcaagc tgtgcangcg gttcatcttg antctgacgc ttctctantt tgtctcaacc 180
atgctctgag cacagaaaaa gangaagtga tgggtctgtg tataggggan ttgaatgatg 240
acataaggag tgactccnaa ttacatata ctggaacgga aatgcgcaca gtccnagaaa 300
agatggatac catcagaatt gttcatatcc attctgtcat catcttgccg cgttctgaca 360
aganaaanga ccgtgttnaa atttctccan ancantgtc 400

```

<210> 33
 <211> 400
 <212> DNA
 <213> Mus musculus

```

<400> 33
gatggcggtc tccacaggag ttaaagttcc tcgtaatttt cgcttgttg aagaacttga 60
agaaggacaa aaaggagtag gtgatggtac tgttagctgg ggccttgaag atgatgaaga 120
catgacactt acaagggtga caggcatgat tattgggcca ccaaggacaa actatgaaaa 180
cagaatatat agcctgaaa tagaatgtgg atctaaatac ccagaagctc ctccatcagt 240
tagatttcta acaaaaatta atatgaatgg gatcaataat tccagtggaa tgggtggatgc 300
acggagcata ccagtattag caaaatggca aaattcctat agcattaaag tcatacttca 360

```

aganctaaga ctcttatg . gtccaaagaa aatatgaanc 400

<210> 34
 <211> 400
 <212> DNA
 <213> Mus musculus

<400> 34
 acggctgcga gaagacgaca gaagggggag ctctgcgtc ccagggactc cagtacacca 60
 ccattggcga ttctgagcgt ctctcgcccc ccggctgctg gttagcctgc accagcttct 120
 cgcgcaccaa aaagggaatt ctctgtttg ctganattat actgtgcctg gtgatcttga 180
 ttgcttcag tgcattaca acatcgccct actcctcctc gtcggtgatt gagatgatct 240
 gtgctgctgt cttacttgtc ttctacacgt gtgacctgca ctccaagata tcattcatca 300
 actggccttg gactgacttc ttcagatccc tcataacaac catcctgtac ctgatcacct 360
 ccattgttgt ccttgtagaa agaagaagca gctccagagt 400

<210> 35
 <211> 400
 <212> DNA
 <213> Mus musculus

<400> 35
 tacggctgcg agaagacgac agaaggggct ccgccttctt cccacacctc ggctgcgggc 60
 gacaccggtc ttgacgtcc gggaccccggt gcgacgttta gcccggtgca cgatgcccc 120
 gaaaaagggg ggagatggaa ttaaacccgc tccaattatt ggaagatttg gaacctcact 180
 gaaaattggt atcgttggat tgccaaatgt tgggaaatct accttcttca atgtattaac 240
 caatagtcag gcttcancag aaaacttccc attctgcact attgatccta atgagagcag 300
 agtgcctgtg ccagatgaga gggttcgactt tcttgccag taccataaac cancaagcaa 360
 gatcctgctt tcctaaatgt aatggatatt gctggccttg 400

<210> 36
 <211> 400
 <212> DNA
 <213> Mus musculus

<400> 36
 acggctgcga gaagacgacn gaaggggtcc ttttctctgc gggcggttcc agtccctctt 60
 gagttgggct gtcgtcnant cgcggcgana ccggcgaccc gcagccatga cagaagctga 120
 tgtnaatccg aangcctatc ccctcncana tccccacctc accaagaanc tgctggacct 180
 tggttcnagc tctgttntct acaancagct tcnгааagga ccnatgaacc nccaaaaccc 240
 tcaacagaag catctctgan ttcatgtga tgcngcanan ctgaaccttg gagatctcct 300
 gcacctccct ctctgtcga aaaaaanaat gtccctacct nttgtncnn ccaacaggct 360
 ttangaaag cctntgggtt tccagccntc atccctgtct 400

<210> 37
 <211> 394
 <212> DNA
 <213> Mus musculus

<400> 37
 gaaaggcgcc tagataccgc gatacttgcg ggtatcctgg agctggactc tcctgttctt 60
 gccacatttg actgaagaag gacaacgggg tcaaggtttg aggactgaca gattctanac 120
 ccaggcttcc tcnccccca naaagccacc anacctcaca tggaaacaaa ggccccatgc 180
 ccancgcgcg tccccctaga gganangaaa ttcctgtgtc ttgtnggcgt cactggcagc 240
 gtggccgctc tgaagctgcc tctcctgnta tctannctgt tggacgttcc tngatgtgg 300
 aagcgccgtt ctgacccggt tctccacatt gacctgcgga tgtgggctga cctcatgcta 360
 gtggtcccc tcnatgcna cactctgngg aagt 394

<210> 38
 <211> 315
 <212> DNA
 <213> Mus musculus

<400> 38

```

acggctgcga gaagacg. . gaangggggc aggatggctc ggtggtacag tt agatca 60
anangcatac ccactgagc aagctgatga agcctactgt gagaggcagg gcttgcnat 120
nangcagatt cgattccggt ttgatggaca accaatcaat gaaacagaca ctccacccca 180
nctggagatg gaggatgagg acacnttga tgtattccag cagcagacag gaggatcanc 240
ctcccggagg agcgtcccac acccaacogt tgtcctgacc tgtgctattg agcagtgacc 300
atgctnccac accca 315

```

<210> 39
 <211> 400
 <212> DNA
 <213> Mus musculus

```

<400> 39
acggctgcga gaagacgaca gaaggggaga gcatcatggc gctaagcggc cgactggcat 60
tggccgcgct cagactgtgg ggtccgggag gatgtatctg anctgcttcg gaccananc 120
tgacgtaggg tgggtgcat ggtgggggcc ggcacanca caccagtggt catcccgac 180
ttcagatccc caggggagcg gcctctacag caaccttcag cagtatgaca tcccgatccc 240
tgaagccatc ttgaaacttg gctttttctt tcacaacccc aagccctttt tcatgttggc 300
caaggactgt nccctgngca ctacaggcca atgtcactca ctacttctg aagctcctcc 360
acgacangga ccgcttctgc gctctntaca canaaatcga 400

```

<210> 40
 <211> 349
 <212> DNA
 <213> Mus musculus

```

<400> 40
ggccatcaan cgtgctgttg gtcattggccg actcctcaaa gatggacccc tctaccgcct 60
caacactaag gcagccagtg ctgaaggga agagggctgt gagtcgctct cgtgtttacc 120
tcccgtgtca ctgcttcac acganaaaga caagccggtt gctgaaccaa tcccatctg 180
tagcttctgt cttgttaca aagaacagaa ccgggagaa cagcccgagg aactcgtctc 240
ctgcgaggac tgcggaata ncggatcacc atcgtgttta aagttctccc canagctnac 300
agtgagagtg aagccttaac gtggcttnca ttgantgtaa aacatcact 349

```

<210> 41
 <211> 397
 <212> DNA
 <213> Mus musculus

```

<400> 41
cccacgcgtc cgtatctctc cagggccacc aagcacctct gaagagccat gttccaagct 60
gccggagccg cccaggccac cccctctcat gaagccaaag gcagcagtg cagcagcacg 120
gtacagcgtt ctaagtctct tagcttgcgg gctcaggtga aggagacctg tgcagcctgc 180
cagaagactg tgtacccgat ggagcggctg gtggcagaca agctcatttt ccacaactct 240
tgtttctgtt gaaaacactg ccacacaaaa ctacgcctgg gcagtatatg tgcaatgcac 300
ggtgaatttt actgcagacc tcactttcag cagctgttta agagtaaagg caactacgat 360
gaagggtttc gaaaacagca caaggactct gggccac 397

```

<210> 42
 <211> 158
 <212> PRT
 <213> Mus musculus

```

<400> 42
Met Phe Gln Ala Ala Gly Ala Ala Gln Ala Thr Pro Ser His Glu Ala
  1                               5                               10                               15

Lys Gly Ser Ser Gly Ser Ser Thr Val Gln Arg Ser Lys Ser Phe Ser
                20                               25                               30

Leu Arg Ala Gln Val Lys Glu Thr Cys Ala Ala Cys Gln Lys Thr Val
  35                               40                               45

Tyr Pro Met Glu Arg Leu Val Ala Asp Lys Leu Ile Phe His Asn Ser

```

Page 15 of 56

35					40					45					
Glu	Leu	Ser	Arg	Thr	Ser	Thr	Ser	Gly	Gly	Asn	Pro	Ser	Phe	Val	Pro
	50					55					60				
Leu	Ser	Lys	Phe	Met	Asn	Thr	Gln	Tyr	Phe	Gly	Thr	Ile	Gly	Leu	Gly
	65				70					75					80
Thr	Pro	Pro	Gln	Asn	Phe	Thr	Val	Val	Phe	Asp	Thr	Gly	Ser	Ser	Asn
				85					90					95	
Leu	Trp	Val	Pro	Ser	Thr	Arg	Cys	His	Phe	Phe	Ser	Leu	Ala	Cys	Trp
			100					105					110		
Phe	His	His	Arg	Phe	Asn	Pro	Lys	Ala	Ser	Ser	Ser	Phe	Arg	Pro	Asn
		115					120					125			
Gly	Thr	Lys	Phe	Ala	Ile	Gln	Tyr	Gly	Thr	Gly	Arg	Leu	Ser	Gly	Ile
	130					135					140				
Leu	Ser	Gln	Asp	Asn	Leu	Thr	Ile	Gly	Gly	Ile	His	Asp	Ala	Phe	Val
	145				150					155					160
Thr	Phe	Gly	Glu	Ala	Leu	Trp	Glu	Pro	Ser	Leu	Ile	Phe	Ala	Leu	Ala
				165					170					175	
His	Phe	Asp	Gly	Ile	Leu	Gly	Leu	Gly	Phe	Pro	Thr	Leu	Ala	Val	Gly
		180						185					190		
Gly	Val	Gln	Pro	Pro	Leu	Asp	Ala	Met	Val	Glu	Gln	Gly	Leu	Leu	Glu
		195					200					205			
Lys	Pro	Val	Phe	Ser	Phe	Tyr	Leu	Asn	Arg	Asp	Ser	Glu	Gly	Ser	Asp
	210					215					220				
Gly	Gly	Glu	Leu	Val	Leu	Gly	Gly	Ser	Asp	Pro	Ala	His	Tyr	Val	Pro
	225				230					235					240
Pro	Leu	Thr	Phe	Ile	Pro	Val	Thr	Ile	Pro	Ala	Tyr	Trp	Gln	Val	His
				245					250					255	
Met	Glu	Ser	Val	Lys	Val	Gly	Thr	Gly	Leu	Ser	Leu	Cys	Ala	Gln	Gly
			260					265					270		
Cys	Ser	Ala	Ile	Leu	Asp	Thr	Gly	Thr	Ser	Leu	Ile	Thr	Gly	Pro	Ser
		275					280					285			
Glu	Glu	Ile	Arg	Ala	Leu	Asn	Lys	Ala	Ile	Gly	Gly	Tyr	Pro	Phe	Leu
	290					295					300				
Asn	Gly	Gln	Tyr	Phe	Ile	Gln	Cys	Ser	Lys	Thr	Pro	Thr	Leu	Pro	Pro
	305				310					315					320
Val	Ser	Phe	His	Leu	Gly	Gly	Val	Trp	Phe	Asn	Leu	Thr	Gly	Gln	Asp
				325					330					335	
Tyr	Val	Ile	Lys	Ile	Leu	Gln	Ser	Asp	Val	Gly	Leu	Cys	Leu	Leu	Gly
		340						345					350		
Phe	Gln	Ala	Leu	Asp	Ile	Pro	Lys	Pro	Ala	Gly	Pro	Leu	Trp	Ile	Leu
		355					360					365			
Gly	Asp	Val	Phe	Leu	Gly	Pro	Tyr	Val	Ala	Val	Phe	Asp	Arg	Gly	Asp
	370					375					380				

Lys Asn Val Gly Pro Arg Val Gly Leu Ala Arg Ala Gln Ser Arg Ser
385 390 395 400

Thr Asp Arg Ala Glu Arg Arg Thr Thr Gln Ala Gln Phe Phe Lys Arg
405 410 415

Arg Pro Gly

<210> 45
<211> 3338
<212> DNA
<213> Mus musculus

<400> 45
cggtgacgca ggaccaggac tcgcgcgtcc agcggagaag caggagaagc cggcgacctt 60
gcgctctcag cctgacccct gtcttgccgg cctgaacatt cgcagctgga gagatggcgt 120
tcgtgaagag tggatgggta ctccggcaga gcaccattct gaaacgctgg aagaagaatt 180
ggttcgacct gtggtcagac ggtcacctga tctactacga tgatcagact cggcagagca 240
tagaggataa ggtccacatg cccgtggact gcatcaatat ccgcacgggg catgagtggc 300
gggacatcca gcctccagat gggaagccca gagactgtct gctgcagatc gtttgccgag 360
acgggaagac catcagttct tgtgcagaga gcacagacga ttgcctggca tggaaagttt 420
cactgcagga ttccagaaca aacacagctt acgttggttc agcaatcctg tctgaagaga 480
ctgcagtgcc cgcgtccccc cctccctacg caacctatgc tacaccnacc cctgaggtct 540
acggctatgg tccatacagc ggcgcatacc ccgcaggaac tcaagttgtc tatgccgcca 600
acggggcaggc atatgcagtg ccataccagt acccgatgc aggaagttat ggacaacagc 660
ctgccaacca agtcatcatc cgcgagcggc accgagacaa tgacagtggc ctggctcttg 720
gcatgctcgc cggggcagcc accggcatgg cctggggctc tctgttcttg gtcttctaga 780
gccttcaaca ttttctgtgc atagcttctg ttagtctctg gtgcagtaat ttgatttgca 840
gggcatctct gtttctgtga agtgtcttct ataataattt aaatagttct tttgaaggtg 900
gtaatctaat aattgtgact gacctgcagt gtaccacaaa gaaagcccga ggtatgctgt 960
gagtgcagac ctgagtcctt ccgggtacta gcttgaccca agtcttctct agggactttt 1020
ggatggcttt atgtaaacac acccagttaa atgggcaatt tccgtccagt taggtgcagt 1080
gttgaattaa gggatggctt tccttgctat gccaaatacta atactgctga tggagggaaga 1140
tgtgtgcaag tgtggtgagg agagtcacag cttctttaac tgtggattct ctctcagacc 1200
cctgctgcgt gttaccctag gagctgtggg ctggtggctc ctgcaagact atggtgtgag 1260
gaccctgtaa cgtacctctt ggagcactta ggtaccctct gaagctccta ggtatcacca 1320
gcaggattgg ctgctcagga tgcagagggc caccctctcc ctttaaaaat tacgctccag 1380
taatctgccc agttttatct tcttgttatt cttctgtttg ctttctctgg ggatgattgg 1440
cattagtctg gagttaggaa ttgattcgag tgcgggtggg tggaggcatg caggagagctg 1500
tccagcgacc tgctctcagt gtttgtttta ggtatattga ttgccagctc aggcgcgaga 1560
gagcctatag agactatttt tctacttgta aagaaagtat ggtgagggga attgaggaga 1620
gccttgtttg aatgttctct cctcaggcct cctggggccc acactgcgtg gtccctggga 1680
ggcttttctt cagcagtgga agggaggccc cgtggctgtc cagagctctc aggtttgagt 1740
gagaaaatgg attgggtaga gcatctcagg gatttgttct aatccctcat gttatgggga 1800
tccagccgtg ttctcagtc agaccgctc acctcagaag agcttaaaac atttctggtc 1860
cccaaagtgt tggcactctg agaagctcac aatctggctt tctaacgaaa atttctgatt 1920
ctaaaattag agaatacatg ttccacgcat ttaaaattta tgttctttca tgttttaaag 1980
ctcccaatc cagctttgtg actggcatat tttagtttca aacagtaccc cggcacaaaag 2040
gtgggatggc acagtgaagg cccccgccc tctactttgc atagtcttct ttctccaggg 2100
tgctcccagg aagcattcat tctgactttg ctgagcccag tgcatgcgtg ctgcttggcc 2160
gcccgtctgc tgggtagctc tttcttggtc agatcaagtc ttcaacagat ctccagtga 2220
gacagtggcc aagtagatga ggtggtgccc atagtgtctt ctgcatactc cttggggacc 2280
tggtgacacc tgcccatctc cagctgacat ttgttttctt gtcactctct atagatggga 2340
tatgtgacaa catggtagcg acgcccgttca gtgtcgcttt aataagcatg atgctgattt 2400
tacatcctgt gctgtatgac tgccatttgc tcacagtgtc accattgcta aagctccgtg 2460
ctttacttac aaaacactaa aaccagtggg tagtgtttca cagtgatttt aattttagag 2520
ttagttactg gcattcctaa agccatagag tactgagtca catccctgaa gtacttttga 2580
aacagaattg tctcctactg tcccatgggt gtgccctgct tgtctcctgg ccccaatggg 2640
gctagctgta ccaggcagcc atagttgagc ctgacatctc ctgtcaccag tttgacttga 2700
ttatatcccc agaattgaa acattcttgg gcatctcagt tccctagccc tgactctcat 2760
agacgccacc ctttctgaggt cttttgctggc gtcacttgta cctcagtgag tcttcgagatt 2820

```

cttgagtttag aggggac ttgtccagca ttgaggaaca tgtctctctc ac agactt 2880
aaatgatgat gcagggctgg aagaggctgg ctgctgacac tgcacgtgg ctgatgtcat 2940
tgctctccta gtctttgat ttaagaacct tccatattga aggcctgagg ctccctcaga 3000
tcgtcccttg ccaagaaggc ctggcttagg tcattagtgc ccacagtagc cttctggagt 3060
gtagcaagtt cctgcgtttg agacagaatg gttcagattt attttctaca tctgttggtg 3120
accccatgca cctctcatt ttgccttcca gtctacgtag atgaaagatg aaaggcagag 3180
gatgcagaca gtcttctttg tgattgcttc tgttattctg ttgcatctac cgagcccgtt 3240
ttctccctgt ctgtgcatac agtatgttta taagtgaact tgttaaaata ttaaatgatc 3300
actaacccta aaaaaaaaaa aaaaaaaaaa aaaaaaaa 3338

```

<210> 46
 <211> 221
 <212> PRT
 <213> Mus musculus

```

<400> 46
Met Ala Phe Val Lys Ser Gly Trp Leu Leu Arg Gln Ser Thr Ile Leu
  1          5          10          15

Lys Arg Trp Lys Lys Asn Trp Phe Asp Leu Trp Ser Asp Gly His Leu
  20          25          30

Ile Tyr Tyr Asp Asp Gln Thr Arg Gln Ser Ile Glu Asp Lys Val His
  35          40          45

Met Pro Val Asp Cys Ile Asn Ile Arg Thr Gly His Glu Cys Arg Asp
  50          55          60

Ile Gln Pro Pro Asp Gly Lys Pro Arg Asp Cys Leu Leu Gln Ile Val
  65          70          75          80

Cys Arg Asp Gly Lys Thr Ile Ser Leu Cys Ala Glu Ser Thr Asp Asp
  85          90          95

Cys Leu Ala Trp Lys Phe Thr Leu Gln Asp Ser Arg Thr Asn Thr Ala
  100         105         110

Tyr Val Gly Ser Ala Ile Leu Ser Glu Glu Thr Ala Val Ala Ala Ser
  115         120         125

Pro Pro Pro Tyr Ala Thr Tyr Ala Thr Xaa Thr Pro Glu Val Tyr Gly
  130         135         140

Tyr Gly Pro Tyr Ser Gly Ala Tyr Pro Ala Gly Thr Gln Val Val Tyr
  145         150         155         160

Ala Ala Asn Gly Gln Ala Tyr Ala Val Pro Tyr Gln Tyr Pro Tyr Ala
  165         170         175

Gly Val Tyr Gly Gln Gln Pro Ala Asn Gln Val Ile Ile Arg Glu Arg
  180         185         190

Tyr Arg Asp Asn Asp Ser Asp Leu Ala Leu Gly Met Leu Ala Gly Ala
  195         200         205

Ala Thr Gly Met Ala Leu Gly Ser Leu Phe Trp Val Phe
  210         215         220

```

<210> 47
 <211> 2396
 <212> DNA
 <213> Mus musculus

<400> 47

```

caccacagag gaccccttca acctaaggaa acacccaggc ttcgatagga ccatgctgca 60
gagggtggcag aaaagggaga tcagcaactt tgaataacct atgtacctca acacactggc 120
cggaaggacc tacaatgact acatgcagta tcccggtgtt cccctgggtcc tcgctgacta 180
cacctcagag atgttgaact tgacgaatcc caagactttt cgggatcttt ctaagccaat 240
gggggctcag accaaggaaa ggaagttaga gtttaccag aggtttaaag atgttgaaaa 300
gattgaagga gacatgaccg tgacgtgcca ctactacacc cactattcct cagccatcat 360
tgctgcttcc tacttggtcc gaatgccacc attcacgcag gccttctgct ccttacaggg 420
cggaagcttt gatgtggctg atagaatgtt ccacagtgtt aagagcacgt gggagtctgc 480
ctccaagag aacatgagcg atgtcaggga gctgacacct gaattcttct acctgcccga 540
gtttttaacc aactgtaatg cagtggagtt tggctgcatg caggatggaa cgacactggg 600
ggatgtgcag cttcctccct gggctgacgg ggatccgagg aaattcatca gcttgacag 660
acaggctctg gaaagtgact tcgtcagcag caacctccac cactggatag acctaatatt 720
tgggtataag cagcaggggc cggctgctgt agaggcagtg aacactttcc accctactt 780
ctacggtgat agaatagacc tgggcagcat cactgaccgg ctgatcaaga gcaccatcct 840
gggcttcac agcaactttg gacaggtgcc caagcagatc ttcactaaac cccaccatc 900
cagaacacac acagggaaaa acccagggcc tggaaaggat gcttccaccc ctgtaggcct 960
cccaggccac tcacagtcct tcctccacag cctgccagca ctgagaccct ctcaggtcac 1020
agtcaaaagt atgtaccttt tctctctagg gtcggaatcc cccaaagggg ccacgcgcca 1080
catcgctcct actgagaagt caatcctggc agtgggagaag aacaagctgc tgaagcccc 1140
tctctggaac aggaccttca gctggggtt tgatgacttc agttgctgcc tggggagcta 1200
cggctctgac aagatcctga tgacctttga gaacctggct gcctggggtc cctgtctgtg 1260
cgctgtatgc ccttccccca cgatgatcgt cacatccggg gccagcgagc tgggtgtgcat 1320
ctgggagctg agcctgggtc aaggtcgccc gagaggtctg aaactccgac aggccttgta 1380
tggacacact caggcggtca catgtctgac agcctctgtc accttcagcc tcctgggtgag 1440
cggatccag garcgcactt gtatcctgtg ggacctggac cacctctctc gtgtggcctg 1500
cctgcctgtc caccgggaag gcactctcag cattgccatc agtgaatgtc cgggaacct 1560
tgtctcctgt gccggagccc acttgtcctc gtggaatgtc aatggacagc cctgggccc 1620
tattaccaca gcctggggcc cagaaggaac cataactgtc tgctgcatag tagaggggcc 1680
agcgtgggat gcaagccacg tgatcatcac ggggagtaag gacggaatgg ttcggatttg 1740
gaagacagag gacgtgaaga tgcctgttcc caggcaggca gtgatggagg agcctccac 1800
ggagccccca agccccagag gtcacaagtg ggccaagaa cttgccctga gccgagagct 1860
ggatgtcagt gttgctctga gtggcaagcc cagcaaggca agtcctgctg tgacagctct 1920
ggccatcact aggaaccaga gcaagctcct ggttggcgat gagaagggcg aatcttctgc 1980
tgggtctgtg atgggttaga gacaaggagc tggagggacc gacctgaaag cctggagccc 2040
tggggtcggc agcaacaggc tacaggcaca agatgatgtg tagcctgggc tgcttaacca 2100
gagcaagtgt tgggggggct ccactccaca cagttctcaa ggagtccctg atggtttgca 2160
ccgtgtaccc taaacatgtc tgtagtctat gggacttctg taagaaggat cttggttagac 2220
actgatgtg gaaactgacg ctgcatggga aatttctacg tggctcactt caccaaggct 2280
tattgcactg ggaaaagaag aggggtgggg tattggttca tggaaaccac cccactgtct 2340
ttattttatt aaaactccat ttccgaaaa aaaaaaaaa aaaaaaaa 2396

```

<210> 48

<211> 693

<212> PRT

<213> Mus musculus

<400> 48

```

Thr Thr Glu Asp Pro Phe Asn Leu Arg Lys His Pro Gly Phe Asp Arg
  1             5             10            15

```

```

Thr Met Leu Gln Arg Trp Gln Lys Arg Glu Ile Ser Asn Phe Glu Tyr
      20             25            30

```

```

Leu Met Tyr Leu Asn Thr Leu Ala Gly Arg Thr Tyr Asn Asp Tyr Met
      35             40            45

```

```

Gln Tyr Pro Val Phe Pro Trp Val Leu Ala Asp Tyr Thr Ser Glu Met
      50             55            60

```

```

Leu Asn Leu Thr Asn Pro Lys Thr Phe Arg Asp Leu Ser Lys Pro Met
      65             70            75            80

```

```

Gly Ala Gln Thr Lys Glu Arg Lys Leu Lys Phe Thr Gln Arg Phe Lys

```


8. 90
 Asp Val Glu Lys Ile Glu Gly Asp Met Thr Val Gln Cys His Tyr Tyr
 100 105 110
 Thr His Tyr Ser Ser Ala Ile Ile Val Ala Ser Tyr Leu Val Arg Met
 115 120 125
 Pro Pro Phe Thr Gln Ala Phe Cys Ser Leu Gln Gly Gly Ser Phe Asp
 130 135 140
 Val Ala Asp Arg Met Phe His Ser Val Lys Ser Thr Trp Glu Ser Ala
 145 150 155 160
 Ser Lys Glu Asn Met Ser Asp Val Arg Glu Leu Thr Pro Glu Phe Phe
 165 170 175
 Tyr Leu Pro Glu Phe Leu Thr Asn Cys Asn Ala Val Glu Phe Gly Cys
 180 185 190
 Met Gln Asp Gly Thr Thr Leu Gly Asp Val Gln Leu Pro Pro Trp Ala
 195 200 205
 Asp Gly Asp Pro Arg Lys Phe Ile Ser Leu His Arg Gln Ala Leu Glu
 210 215 220
 Ser Asp Phe Val Ser Ser Asn Leu His His Trp Ile Asp Leu Ile Phe
 225 230 235 240
 Gly Tyr Lys Gln Gln Gly Pro Ala Ala Val Glu Ala Val Asn Thr Phe
 245 250 255
 His Pro Tyr Phe Tyr Gly Asp Arg Ile Asp Leu Gly Ser Ile Thr Asp
 260 265 270
 Pro Leu Ile Lys Ser Thr Ile Leu Gly Phe Ile Ser Asn Phe Gly Gln
 275 280 285
 Val Pro Lys Gln Ile Phe Thr Lys Pro His Pro Ser Arg Asn Thr Thr
 290 295 300
 Gly Lys Asn Pro Gly Pro Gly Lys Asp Ala Ser Thr Pro Val Gly Leu
 305 310 315 320
 Pro Gly His Ser Gln Ser Phe Leu His Ser Leu Pro Ala Leu Arg Pro
 325 330 335
 Ser Gln Val Thr Val Lys Asp Met Tyr Leu Phe Ser Leu Gly Ser Glu
 340 345 350
 Ser Pro Lys Gly Ala Ile Gly His Ile Val Pro Thr Glu Lys Ser Ile
 355 360 365
 Leu Ala Val Glu Lys Asn Lys Leu Leu Met Pro Pro Leu Trp Asn Arg
 370 375 380
 Thr Phe Ser Trp Gly Phe Asp Asp Phe Ser Cys Cys Leu Gly Ser Tyr
 385 390 395 400
 Gly Ser Asp Lys Ile Leu Met Thr Phe Glu Asn Leu Ala Ala Trp Gly
 405 410 415
 Pro Cys Leu Cys Ala Val Cys Pro Ser Pro Thr Met Ile Val Thr Ser
 420 425 430

Gly Ala Ser Ala Val Val Cys Ile Trp Glu Leu Ser Leu Val Lys Gly
 435 440 445
 Arg Pro Arg Gly Leu Lys Leu Arg Gln Ala Leu Tyr Gly His Thr Gln
 450 455 460
 Ala Val Thr Cys Leu Thr Ala Ser Val Thr Phe Ser Leu Leu Val Ser
 465 470 475 480
 Gly Ser Gln Asp Arg Thr Cys Ile Leu Trp Asp Leu Asp His Leu Ser
 485 490 495
 Arg Val Ala Cys Leu Pro Val His Arg Glu Gly Ile Ser Ala Ile Ala
 500 505 510
 Ile Ser Asp Val Ser Gly Thr Ile Val Ser Cys Ala Gly Ala His Leu
 515 520 525
 Ser Leu Trp Asn Val Asn Gly Gln Pro Leu Ala Ser Ile Thr Thr Ala
 530 535 540
 Trp Gly Pro Glu Gly Thr Ile Thr Cys Cys Cys Ile Val Glu Gly Pro
 545 550 555 560
 Ala Trp Asp Ala Ser His Val Ile Ile Thr Gly Ser Lys Asp Gly Met
 565 570 575
 Val Arg Ile Trp Lys Thr Glu Asp Val Lys Met Pro Val Pro Arg Gln
 580 585 590
 Ala Val Met Glu Glu Pro Ser Thr Glu Pro Leu Ser Pro Arg Gly His
 595 600 605
 Lys Trp Ala Lys Asn Leu Ala Leu Ser Arg Glu Leu Asp Val Ser Val
 610 615 620
 Ala Leu Ser Gly Lys Pro Ser Lys Ala Ser Pro Ala Val Thr Ala Leu
 625 630 635 640
 Ala Ile Thr Arg Asn Gln Ser Lys Leu Leu Val Gly Asp Glu Lys Gly
 645 650 655
 Glu Ser Ser Ala Gly Leu Leu Met Gly Arg Arg Gln Gly Ala Gly Gly
 660 665 670
 Thr Asp Leu Lys Ala Trp Ser Pro Gly Val Gly Ser Asn Arg Leu Gln
 675 680 685
 Ala Gln Asp Asp Val
 690

<210> 49
 <211> 2363
 <212> DNA
 <213> Mus musculus

<400> 49
 cagaggcagg caggacaggc actgcasgca ctgagagcct tgcagtcaga agggaagcac 60
 asacaggaag aggcagtgag gagaaaaagga gcasaactga gacaagcaat ggagatccct 120
 gaggagtgtc ggccaacggc tgaagtgtcc ctaaaagaca tcaactgaaat aatggagaga 180
 catctcagtc acatggaacg gacctgtct caccgtcaaa agctctcaga tggagacctg 240
 gtaagatggg catctggagg gctggtcctg cagggaattt ataagaccaa ccaccaaga 300

```

agcctgattc agaagag. agagctactc agtgtcccta agcagttctc ac gctggc 360
ccagtacatg gcacagagat aaaaacaaag gaattttcat cttttcaaga acaagccatg 420
tttacacaga ctatagagag ggtgggtttc agttcaaccc ccttagttaa aggagaaggt 480
tggggactta gtctagaagc tggtaggggt cacaacaaac agacagaatc tgaagataac 540
taccagtccc attcaaagca aacttatttt tgctcagcca ggttcagcta catccccattg 600
gccacctgcc attttcacat caatgatctt gaactctccc aggctgctct ccagggaacta 660
aaaagtattg aagaaatcct ggagcagact acacaccacc gagatggact acccttactg 720
aggcacaggg ctaaaaaactt ttccacaggg tttggctctc atgctaacca aggccctgtg 780
cacctggggg gaactctactg ctggaaagcc atttcagaag gtttcaaaag tgagcacttg 840
gctgatgtaa agcagcaagc agaagagtct ttgaatattt acattatggg cagttatagt 900
ggctttggag ttaaagttagg tgcgagtgtg aatataacaa attcaaaatc aaaaacagca 960
ttttacagta aaactcatct aaactcgcaa accaaggtag aactatctgt agccaagata 1020
ggtggaccag cagaagcaga tggaaatgcc cagtggacag ctggccttgt agctagcaat 1080
caaacctggt ctgttattga taggaaactg cagtgggtac ctatttggga cattatcctg 1140
tccagtcaca gaactgaatt taagaatgct cttcaactgg ctaactgcct caaagaccac 1200
tacactgctc tgactgaact agctgcccag attcaagaag gggaagaatt tctgactgct 1260
agaaaagaag ctaagctttt ctaaagaat gtgaaaggct gggagggtttc tgatcctgaa 1320
gaacagctta ggaagttagt agattttatg caaacattga gtcaaaaaat aaaaagttagt 1380
gacatttggg ttaacacatg cctcatagat tgggatctgc aaaattttct aataaacatt 1440
gtcaacttct gcaaaaaatc acccacttat aaaactcact ttattaaatc tcagtgtgtg 1500
agccttctag aacctcatgt ctacaaagtg acaactttc ctcaggcaca atccatcata 1560
cagtggatca atcagtcaga gtcagaggaa gaactagtca aaattacctc attttctgaa 1620
ttcattaaaca ccttaaagaa aaccacaaaa tacctaattg aagagagttt caaaactgag 1680
ccccagaaa gagtgggaaga agcaaagaga atggctacat atgaagtcac cacagctctc 1740
agctccttct tgaagtacct cagagaaaca cagcagccag acatgcagct gttgctactc 1800
tccattgcta ctggtgtagg ctatcagttg gtaaacagta tttttcagca tcttctgggg 1860
tgtgatgagt taaacttcct cttggatcaa atggaaaata acgaacataa ataccaagaa 1920
ctgaaaaata tttgcaatta cagagccag gcattcttgg tgctcacagc cctaagagcc 1980
acagttgaaa tcacagatgt ttctacagaa gagaaaggac aacgtttgac attaatataa 2040
caacatatgg ggtcactggt gtctgaaagaa gttgcacatg tcttcacaaa acatggagaa 2100
catcatgact gggaaaggct ggagaatgat ttgagattac tcattgaggg ggactataaa 2160
gccaccaccc attccttaca aatggatgaa gtaaaaaaac aattgcaaag tttntgcat 2220
gaaaagaaac agacttataa acaacaagggt aatgaaaaca gaacaaaaga aatgatagaa 2280
aatggacatt tcctggactt actccaacgt ttaggcctag acaattacta tccaaaaaaa 2340
aaaaaaaaa aaaaaaaaaa aaa 2363

```

<210> 50
 <211> 778
 <212> PRT
 <213> Mus musculus

<400> 50
 Gln Arg Gln Ala Gly Gln Ala Leu Xaa Ala Leu Arg Ala Leu Gln Ser
 1 5 10 15
 Glu Gly Lys His Xaa Gln Glu Glu Val Arg Arg Lys Gly Ala Xaa
 20 25 30
 Leu Arg Gln Ala Met Glu Ile Pro Glu Glu Cys Trp Pro Thr Ala Glu
 35 40 45
 Val Ser Leu Lys Asp Ile Thr Glu Ile Met Glu Arg His Leu Ser His
 50 55 60
 Met Glu Arg Thr Leu Ser His Ser Gln Lys Leu Ser Asp Gly Asp Leu
 65 70 75 80
 Val Arg Trp Ala Ser Gly Gly Leu Val Leu Gln Gly Ile Tyr Lys Thr
 85 90 95
 Asn His Pro Arg Ser Leu Ile Gln Lys Arg Glu Glu Leu Leu Ser Val
 100 105 110
 Pro Lys Gln Phe Ser Leu Ala Gly Pro Val His Gly Thr Glu Ile Lys

115	120	125
Thr Lys Glu Phe Ser Ser Phe Gln Glu Gln Ala Met Phe Thr Gln Thr		
130	135	140
Ile Glu Arg Val Gly Phe Ser Ser Thr Pro Leu Val Lys Gly Glu Gly		
145	150	155 160
Trp Gly Leu Ser Leu Glu Ala Gly Met Gly His Asn Lys Gln Thr Glu		
	165 170	175
Ser Glu Asp Asn Tyr Gln Ser His Ser Lys Gln Thr Tyr Phe Cys Ser		
	180 185	190
Ala Arg Phe Ser Tyr Ile Pro Leu Ala Thr Cys His Phe His Ile Asn		
	195 200	205
Asp Leu Glu Leu Ser Gln Ala Ala Leu Gln Glu Leu Lys Ser Ile Glu		
	210 215	220
Glu Ile Leu Glu Gln Thr Thr His His Arg Asp Gly Leu Pro Leu Leu		
225	230	235 240
Arg His Arg Ala Lys Asn Phe Phe His Arg Phe Gly Ser His Ala Asn		
	245	250 255
Gln Gly Pro Val His Leu Gly Gly Ile Tyr Cys Trp Lys Ala Ile Ser		
	260	265 270
Glu Gly Phe Lys Ser Glu His Leu Ala Asp Val Lys Gln Gln Ala Glu		
	275	280 285
Glu Ser Leu Asn Ile Tyr Ile Met Gly Ser Tyr Ser Gly Phe Gly Val		
	290 295	300
Lys Val Gly Ala Ser Val Asn Ile Thr Asn Ser Lys Ser Lys Thr Ala		
305	310	315 320
Phe Tyr Ser Lys Thr His Leu Asn Ser Gln Thr Lys Val Gln Leu Ser		
	325	330 335
Val Ala Lys Ile Gly Gly Pro Ala Glu Ala Asp Gly Ile Ala Gln Trp		
	340	345 350
Thr Ala Gly Leu Val Ala Ser Asn Gln Thr Trp Ser Val Ile Asp Arg		
	355	360 365
Lys Leu Gln Leu Val Pro Ile Trp Asp Ile Ile Leu Ser Ser His Arg		
	370	375 380
Thr Glu Phe Lys Asn Ala Leu Gln Leu Ala Asn Cys Leu Lys Asp His		
385	390	395 400
Tyr Thr Ala Leu Thr Glu Leu Ala Ala Gln Ile Gln Glu Gly Glu Glu		
	405	410 415
Phe Leu Thr Ala Arg Lys Glu Ala Lys Leu Phe Leu Lys Asn Val Lys		
	420	425 430
Gly Trp Glu Val Ser Asp Pro Glu Glu Gln Leu Arg Lys Leu Val Asp		
	435	440 445
Phe Met Gln Thr Leu Ser Gln Lys Ile Lys Ser Tyr Asp Ile Trp Ile		
	450	455 460

Asn Thr Cys Leu Ile Asp Trp Asp Leu Gln Asn Phe Leu Ile Asn Ile
 465 470 475 480
 Val Asn Phe Cys Lys Asn Ser Pro Thr Tyr Lys Thr His Phe Ile Lys
 485 490 495
 Ser Gln Leu Cys Ser Leu Leu Glu Pro His Val Tyr Lys Val Thr Asn
 500 505 510
 Phe Pro Gln Ala Gln Ser Ile Ile Gln Trp Ile Asn Gln Ser Glu Ser
 515 520 525
 Glu Glu Glu Leu Val Lys Ile Thr Ser Phe Ser Glu Phe Ile Asn Thr
 530 535 540
 Leu Lys Lys Thr His Lys Tyr Leu Met Glu Glu Ser Phe Lys Thr Glu
 545 550 555 560
 Pro Pro Glu Arg Val Glu Glu Ala Lys Arg Met Ala Thr Tyr Glu Val
 565 570 575
 Thr Thr Ala Leu Ser Ser Phe Leu Lys Tyr Leu Arg Glu Thr Gln Gln
 580 585 590
 Pro Asp Met Gln Leu Leu Leu Leu Ser Ile Ala Thr Gly Val Gly Tyr
 595 600 605
 Gln Leu Val Asn Ser Ile Phe Gln His Leu Leu Gly Cys Asp Glu Leu
 610 615 620
 Asn Phe Leu Leu Asp Gln Met Glu Asn Asn Glu His Lys Tyr Gln Glu
 625 630 635 640
 Leu Lys Asn Ile Cys Asn Tyr Arg Ala Gln Ala Phe Leu Val Leu Thr
 645 650 655
 Ala Leu Arg Ala Thr Val Glu Ile Thr Asp Val Ser Thr Glu Glu Lys
 660 665 670
 Gly Gln Arg Leu Thr Leu Ile Gln Gln His Met Gly Ser Leu Leu Ser
 675 680 685
 Glu Glu Val Ala His Val Leu Thr Lys His Gly Glu His His Asp Trp
 690 695 700
 Glu Arg Leu Glu Asn Asp Leu Arg Leu Leu Ile Glu Gly Asp Tyr Lys
 705 710 715 720
 Ala Thr Thr His Ser Leu Gln Met Asp Glu Val Lys Lys Gln Leu Gln
 725 730 735
 Ser Xaa Cys His Glu Lys Lys Gln Thr Tyr Lys Gln Gln Gly Asn Glu
 740 745 750
 Asn Arg Thr Lys Glu Met Ile Glu Asn Gly His Phe Leu Asp Leu Leu
 755 760 765
 Gln Arg Leu Gly Leu Asp Asn Tyr Tyr Pro
 770 775

<210> 51
 <211> 1712

<212> DNA

<213> Mus musculus

<400> 51

```

tcttcgaaag cccggctgag gggaatcctg gacaggggaa tccctggacgt ggagatcgtg 60
agtcattggt gcttcccag agcgtgatga gatccacaag gacgttcaga actactatgg 120
gaatgtactg aagacatctg cagacctcca gactaatgct tgtgtcacgc gagccaagcc 180
ggccccagc tacatccggg aaagtctgca gaatgtacac gaagacgtta gttcgaggta 240
ttatggctgt ggtctgactg ttcctgagcg gctggaaaac tgccgaattt tggatctggg 300
tagtgggagt ggcagggtt gctatgtgct tagccagctg gttgggtgaga agggacatgt 360
caccggaata gacatgactg aggtccaggt cgaagtggct aaaacctatc ttgaacacca 420
catggaaaaa tttggtttcc aggcacccaa tgtgactttt ctccacggcc gcatcgaga 480
gttggcagag gctgggatcc agagtggag ctatggtatt gtcatatcca actgtgttat 540
caaccttggt cctgataaac aacaagtcct ccaggaggtc tatcgagtgc tgaagcacgg 600
cggggagctc tatttcagtg acgtctatgc cagccttgaa gtgccagaag acatcaagtc 660
gcacaaagt tttatggggg aatgcctggg aggcgctctg tactggaagg atcttgccat 720
cattgcccac aagattgggt tctgcctcc acgtttggtc actgccgata tcattactgt 780
tgaaaacaag gagctcgaag ggggtcttgg tgactgtcgc tttgtgtctg ccacatttcg 840
cctcttcaaa ctccctaaga cagagccagc cgaagatgc cgagttgttt acaatggagg 900
aatcaaggga catgaaaagg aactaatttt cgatgcaa atcacattca aggaaggcga 960
agctgttgca gtggatgagg agacggcagc tgcctgaag aactcacgtt ttgctccgga 1020
ttttctctc acacctgtg acgcctcgct gccagctccc caggggccgt tctgagttag 1080
agacaaagt tctaatacaga gatccattca agcttgcaga ggactctgac aagatgaagc 1140
ccagacatgc acctgaaggc acgggaggct gctgtggcaa gaggaaaaac tgctagatct 1200
acagccagcg cggagccac cgggctcaag aggggtggta aaggacagtc acagaggctt 1260
cttagcctgc tcttcgccag tgcacagatt atgtgaaggt ggcaaaagcca ccacaagcta 1320
gacctgtct agaataaga gtgactttta gaggatgtta attgaaggtt cacagcaaat 1380
cgctgtctt tctatttctc tatctcagag ttctgtgccc acctagtggc cagaagttaga 1440
acttggaagc ccaagggtta ctcaaagggc caaaggcatc atcaacgttg tgagaattat 1500
cttctctctg gcctaccaca ggacacctct gggttcttct ctgtggttac caggaagcac 1560
agtacttact aaatttatgc taaccatgac aaaagattgt caactcaaat ttgctaggag 1620
tattctttag gttgctgtct gcaattttt tctctgtaac tgaatgaaaa agaaaaacaa 1680
taaaaaataa atttgacttc gaaaaaaaaa aa 1712

```

<210> 52

<211> 336

<212> PRT

<213> Mus musculus

<400> 52

```

Met Ala Ala Ser Arg Asp Ala Asp Glu Ile His Lys Asp Val Gln Asn
  1             5             10             15

Tyr Tyr Gly Asn Val Leu Lys Thr Ser Ala Asp Leu Gln Thr Asn Ala
      20             25             30

Cys Val Thr Arg Ala Lys Pro Val Pro Ser Tyr Ile Arg Glu Ser Leu
      35             40             45

Gln Asn Val His Glu Asp Val Ser Ser Arg Tyr Tyr Gly Cys Gly Leu
      50             55             60

Thr Val Pro Glu Arg Leu Glu Asn Cys Arg Ile Leu Asp Leu Gly Ser
      65             70             75             80

Gly Ser Gly Arg Asp Cys Tyr Val Leu Ser Gln Leu Val Gly Glu Lys
      85             90             95

Gly His Val Thr Gly Ile Asp Met Thr Glu Val Gln Val Glu Val Ala
      100            105            110

Lys Thr Tyr Leu Glu His His Met Glu Lys Phe Gly Phe Gln Ala Pro
      115            120            125

```

Asn Val Thr Phe Leu is Gly Arg Ile Glu Lys Leu Ala Glu ; Gly
 130 135 140
 Ile Gln Ser Glu Ser Tyr Gly Ile Val Ile Ser Asn Cys Val Ile Asn
 145 150 155 160
 Leu Val Pro Asp Lys Gln Gln Val Leu Gln Glu Val Tyr Arg Val Leu
 165 170 175
 Lys His Gly Gly Glu Leu Tyr Phe Ser Asp Val Tyr Ala Ser Leu Glu
 180 185 190
 Val Pro Glu Asp Ile Lys Ser His Lys Val Leu Trp Gly Glu Cys Leu
 195 200 205
 Gly Gly Ala Leu Tyr Trp Lys Asp Leu Ala Ile Ile Ala Gln Lys Ile
 210 215 220
 Gly Phe Cys Pro Pro Arg Leu Val Thr Ala Asp Ile Ile Thr Val Glu
 225 230 235 240
 Asn Lys Glu Leu Glu Gly Val Leu Gly Asp Cys Arg Phe Val Ser Ala
 245 250 255
 Thr Phe Arg Leu Phe Lys Leu Pro Lys Thr Glu Pro Ala Glu Arg Cys
 260 265 270
 Arg Val Val Tyr Asn Gly Gly Ile Lys Gly His Glu Lys Glu Leu Ile
 275 280 285
 Phe Asp Ala Asn Phe Thr Phe Lys Glu Gly Glu Ala Val Ala Val Asp
 290 295 300
 Glu Glu Thr Ala Ala Val Leu Lys Asn Ser Arg Phe Ala Pro Asp Phe
 305 310 315 320
 Leu Phe Thr Pro Val Asp Ala Ser Leu Pro Ala Pro Gln Gly Pro Phe
 325 330 335

<210> 53

<211> 3599

<212> DNA

<213> Mus musculus

<400> 53

agagacagcg tgatcccggc ctcccacggg gcagctttta ctgtctaggg aagaaatccc 60
 caaagtccat ggagtctgaa gactctgtca agcctcgcta ggaaacctag gagttttaga 120
 gggcacttgg caccggaagc tagccgggta ggcggagcct cacctggatt gatttcacag 180
 ctgcctagac aggctcagac taggtgctgg gcacctggga ggaggaggag acattagcag 240
 caaaggctgt taacagaagt gcctgcctag gcttggaggc aagacgctgc tgttcacagt 300
 gcgagacgga ggtaggagta taatggctgt ccaggtgctg cggcagatgg tctacttcct 360
 actgagctctg tttctctctg tgcaaggtgc acacagtggc agcccccgag aagacttccg 420
 cttctgtggc cagcggaacc agacccaaca gagcaccctc cactatgata aatcttcaga 480
 gcctcacatc tttgtgtgga acacagagga gaccctcaca attcgtgccc ccttcctggc 540
 agccccgat attccccgct tcttcccaga gcctagaggg ctctatcact tctgcctcta 600
 ctggagtgcg cactctggga gactccactt gcgctatggc aagcatgact acctgcttag 660
 tagccaagcc tccagactcc tctgcttcca gaaacaggag cagagcctga agcaggaggc 720
 cccgctgata gccacctctg tcagctcctg gcagattccc cagaacacca gcctgcctgg 780
 ggctccgagc ttcatcttct ccttcacaaa tgccccacac aaggtctccc acaatgcata 840
 tgtggacatg tgtgatctca agaaggaatt gcagcagctt agcaggtacc tgcagcacc 900

```

tcaaaaggct gccaaagc ccaccgcagc gttcatcagc cagcagttac ac jctgga 960
gtcaaaagctg acctctgtga gcttctctggg agacacatta tcctttgagg aggaccgggt 1020
caatgctaca gtgtggaagc tgccaccac agccgggtcta gaggatctgc atatccactc 1080
ccagaaggag gaggagcaga gtgaggtcca ggcatactcg ctgttgcttc cccggggcgt 1140
attccagcag accagaggcc gtcgcccggga tgacgccaag aggtccttg tagtagactt 1200
cagcagccaa gctttgttcc aggacaagaa ttctagccaa gtcttgggtg agaaggtctt 1260
gggtattgtc gtgcagaaca ccaaagtcac caacctctca gatccgggtg tactcacctt 1320
ccagcaccag cctcagccaa aaaatgtgac tctgcagtgc gtgttctggg ttgaagaccc 1380
ggcatcaagc agcacagga gctggagcag tgcaggctgc gagacagtga gcagagacac 1440
acagacatcc tgctgtgca accacctgac ctactttgca gtgctgatgg tgcattccac 1500
agaggtagaa gccactcaca aacactacct cagctctctg toctacgtgg gctgtgtcat 1560
ctctgctctg gcttgtgtct tcactatcgc tgccctacctc tgctccagga ggaagtcacg 1620
tgactacacc atcaaagtc acatgaacct gctgtccgct gtcttctctg tggacgtgag 1680
cttctgtctc agcagagcgt tggcactgac gggctccgaa gcagcctgtc gcaccagtgc 1740
catgttctcg cacttctccc tgcttgcctg cctctctctg atgggctctg agggctacaa 1800
tctctaccga ctggtggtgg aggtctctcg tacttatgtg cccggctatc tgctcaagct 1860
gagcatctg ggctggggtt ttctgtctt cctggtcact ctggtggcgt tgggtgatgt 1920
gaataactac ggcaccatta tcctagctgt gcgcccggact ccggaacctg tcacctaccc 1980
ctctatgtgc tggatccggg actccctggt gagctatgtc accaacctgg gcctcttcag 2040
tctggtgttc ctgttcaacc tggctatgct ggccaccatg gtggtgcaga tcctgcggt 2100
tcgcccgcac agccagaact ggcaccacgt gctgaccctg ctgggctca gcctggctct 2160
tggcctcccc tgggcttgg tcttctttt ctttgccttc ggcaccttc agcttgtcat 2220
cctctaccto ttcagcatca taacttccta ccaaggcttc ctcatcttc tggtgtagt 2280
gtccatgctg ttccaggccc aaggcgccc ctccctctg aagaacaact cagacagcgc 2340
caaactcccc atcagctccg gcagcacctc ctccagccgc atctaagcca ccgccacacc 2400
tcccctccgg gaggacacat gcattggcgtc cgctcacgat gtctgtggcc cagtgtctgtg 2460
cccaccagc ctttgttgg tagtggcata ctagagaagg ccctggtcct tgaaggcgt 2520
gggctgttgc tctgttaggt agatacctag cttgccttgg ggacgactct ggtcctcaaa 2580
gggctcagaa gcacactgcc attctgtttg tggggcgtt tcagtctgga gctaaggcct 2640
tgtcttctg gccacctctg ggtccagctg ttgtgtctgg gtgttgagac ctgcagaccc 2700
aagctggggt tagatctcga aggaggtcga cacatccggc ctgagacaca gctaactgtc 2760
ttgacttgct gctctgtctc tgtggtcacc atgcagatcc cgagggtggc actgggggta 2820
aatgttctg gagaagggtt ggaggcagag caccttagga gctgagcatc tccccagcc 2880
tttctgcaaa cctctctctt cattcccat ccccaacccc tctctgtgt tccccaaacc 2940
ctccacctga agcctggggt cctagaccaa tgctgtgatt tggggtggtg gttcccagca 3000
gttctctggt gccagctatc aacttctgtc tgttgtgtgg gcttggcct ctgactcagg 3060
gcaggtttct gtctgagccc tctctccaag ctgcctcacc ttgtctcgca cctcagagg 3120
acctccatct ctctgaagc ctctccctc tggcaagtac tgggatacag ccaccttctc 3180
aaccagcac tctgaagacc aagacagccc cctctgttga cactggccaa gcttgatctt 3240
tttcttaaga agtgtcttc agatcccgc aggtcgctca gaagacactg ggctgcctag 3300
tgtgaattct gtcctactaa cgtacagtga gcagctctc accccacccc ccgcaaaagc 3360
tctcaccaag tctggagtg tcaggcagg ggctggaat ccaggaggac ttcctgcaaa 3420
aggcagcatt tcactctgac ctacgcttc aggttgggga gaatgttctt tttaaatacc 3480
agttcatttg tcttttgata ttaaagctct ttatagagag tctggaaact gtaggcgatt 3540
gtcgagaaga gaaataaaaa tgagctgtta tctaattgcca tggcaagca gcacaaaaa 3599

```

<210> 54
 <211> 687
 <212> PRT
 <213> Mus musculus

<400> 54
 Met Ala Val Gln Val Leu Arg Gln Met Val Tyr Phe Leu Leu Ser Leu
 1 5 10 15
 Phe Ser Leu Val Gln Gly Ala His Ser Gly Ser Pro Arg Glu Asp Phe
 20 25 30
 Arg Phe Cys Gly Gln Arg Asn Gln Thr Gln Gln Ser Thr Leu His Tyr
 35 40 45
 Asp Gln Ser Ser Glu Pro His Ile Phe Val Trp Asn Thr Glu Glu Thr
 50 55 60

Leu Thr Ile Arg Ala Pro Phe Leu Ala Ala Pro Asp Ile Pro A Phe
 65 70 75 80
 Phe Pro Glu Pro Arg Gly Leu Tyr His Phe Cys Leu Tyr Trp Ser Arg
 85 90 95
 His Thr Gly Arg Leu His Leu Arg Tyr Gly Lys His Asp Tyr Leu Leu
 100 105 110
 Ser Ser Gln Ala Ser Arg Leu Leu Cys Phe Gln Lys Gln Glu Gln Ser
 115 120 125
 Leu Lys Gln Gly Ala Pro Leu Ile Ala Thr Ser Val Ser Ser Trp Gln
 130 135 140
 Ile Pro Gln Asn Thr Ser Leu Pro Gly Ala Pro Ser Phe Ile Phe Ser
 145 150 155 160
 Phe His Asn Ala Pro His Lys Val Ser His Asn Ala Ser Val Asp Met
 165 170 175
 Cys Asp Leu Lys Lys Glu Leu Gln Gln Leu Ser Arg Tyr Leu Gln His
 180 185 190
 Pro Gln Lys Ala Ala Lys Arg Pro Thr Ala Ala Phe Ile Ser Gln Gln
 195 200 205
 Leu Gln Ser Leu Glu Ser Lys Leu Thr Ser Val Ser Phe Leu Gly Asp
 210 215 220
 Thr Leu Ser Phe Glu Glu Asp Arg Val Asn Ala Thr Val Trp Lys Leu
 225 230 235 240
 Pro Pro Thr Ala Gly Leu Glu Asp Leu His Ile His Ser Gln Lys Glu
 245 250 255
 Glu Glu Gln Ser Glu Val Gln Ala Tyr Ser Leu Leu Leu Pro Arg Ala
 260 265 270
 Val Phe Gln Gln Thr Arg Gly Arg Arg Arg Asp Asp Ala Lys Arg Leu
 275 280 285
 Leu Val Val Asp Phe Ser Ser Gln Ala Leu Phe Gln Asp Lys Asn Ser
 290 295 300
 Ser Gln Val Leu Gly Glu Lys Val Leu Gly Ile Val Val Gln Asn Thr
 305 310 315 320
 Lys Val Thr Asn Leu Ser Asp Pro Val Val Leu Thr Phe Gln His Gln
 325 330 335
 Pro Gln Pro Lys Asn Val Thr Leu Gln Cys Val Phe Trp Val Glu Asp
 340 345 350
 Pro Ala Ser Ser Ser Thr Gly Ser Trp Ser Ser Ala Gly Cys Glu Thr
 355 360 365
 Val Ser Arg Asp Thr Gln Thr Ser Cys Leu Cys Asn His Leu Thr Tyr
 370 375 380
 Phe Ala Val Leu Met Val Ser Ser Thr Glu Val Glu Ala Thr His Lys
 385 390 395 400
 His Tyr Leu Thr Leu Leu Ser Tyr Val Gly Cys Val Ile Ser Ala Leu

401 410
 Ala Cys Val Phe Thr Ile Ala Ala Tyr Leu Cys Ser Arg Arg Lys Ser
 420 425 430
 Arg Asp Tyr Thr Ile Lys Val His Met Asn Leu Leu Ser Ala Val Phe
 435 440 445
 Leu Leu Asp Val Ser Phe Leu Leu Ser Glu Pro Val Ala Leu Thr Gly
 450 455 460
 Ser Glu Ala Ala Cys Arg Thr Ser Ala Met Phe Leu His Phe Ser Leu
 465 470 475 480
 Leu Ala Cys Leu Ser Trp Met Gly Leu Glu Gly Tyr Asn Leu Tyr Arg
 485 490 495
 Leu Val Val Glu Val Phe Gly Thr Tyr Val Pro Gly Tyr Leu Leu Lys
 500 505 510
 Leu Ser Ile Val Gly Trp Gly Phe Pro Val Phe Leu Val Thr Leu Val
 515 520 525
 Ala Leu Val Asp Val Asn Asn Tyr Gly Pro Ile Ile Leu Ala Val Arg
 530 535 540
 Arg Thr Pro Glu Arg Val Thr Tyr Pro Ser Met Cys Trp Ile Arg Asp
 545 550 555 560
 Ser Leu Val Ser Tyr Val Thr Asn Leu Gly Leu Phe Ser Leu Val Phe
 565 570 575
 Leu Phe Asn Leu Ala Met Leu Ala Thr Met Val Val Gln Ile Leu Arg
 580 585 590
 Leu Arg Pro His Ser Gln Asn Trp Pro His Val Leu Thr Leu Leu Gly
 595 600 605
 Leu Ser Leu Val Leu Gly Leu Pro Trp Ala Leu Val Phe Phe Ser Phe
 610 615 620
 Ala Ser Gly Thr Phe Gln Leu Val Ile Leu Tyr Leu Phe Ser Ile Ile
 625 630 635 640
 Thr Ser Tyr Gln Gly Phe Leu Ile Phe Leu Trp Tyr Trp Ser Met Arg
 645 650 655
 Phe Gln Ala Gln Gly Gly Pro Ser Pro Leu Lys Asn Asn Ser Asp Ser
 660 665 670
 Ala Lys Leu Pro Ile Ser Ser Gly Ser Thr Ser Ser Ser Arg Ile
 675 680 685

<210> 55

<211> 3590

<212> DNA

<213> Mus musculus

<400> 55

ctcagggaag gccgatcttc cgggtggagg gggaagcggc gtgactggag tggaaatttt 60
 tcccaacaca accttcgga ggcaacatat tggaaggac tcggggaggc cggagtccaa 120
 atggaagtgg ctgaaagaaa cttctcgccc tgctgattct gagccccgcg tcgtgccgcg 180
 cgccctcaat taccatcatg acctgtggtc ttgacagaac attcttcaca atccaaaaga 240

```

aaaagcagac tggttcg; tgtttgacgg ctatgatagc tgcagtgagg ac lagtag 300
cagctccagc tctgaggaga gtgaagaaga agttgctccc ttacctcca atctcccaat 360
catcaagaat aatggacaag tctacacata ccagatggg- aaatctggca tggctacctg 420
tgagatgtgt gggatggtcg gtgtgcgaga tgctttttac tctaaaacga aacgtttctg 480
cagtgtttcc tgttcaagaa gttactcgtc aaactctaag aaggcaagca ttctggccag 540
acttcagggt aagcctccaa caaagaaagc aaaagtcct- caaaaacaac ctttagttgc 600
taaattagct gcctatgccc agtatcaagc taccttgcaa aatcaagcaa agacaaaagc 660
aggcaattct gcaatctctg tggaaaggctt cagctgggg- aactacatca acagcaacag 720
cttcatagca gctccagtg- cctgttttaa gcatgcacct atggggacct gctggggtga 780
tatctcgga- aatgtaagga tagaagttcc caatacagac tgcagtctac ctaccaaagt 840
cttctggatt gctggaatta taaaattagc aggttataat gcccttttga gatatgaagg 900
atttgaaaat gattcttctc tggacttctg gtgcaatata tgtgggtctg atattcatcc 960
agttgggttg tgtgcagcta gtggaiaacc tctcgttccc cctagaactg ttcaacataa 1020
atatacaaac tggaaagctt ttctagtaaa aagacttact ggtgcaaaa cacttccctc 1080
tgaattctca caaaaggttt ctgagagtat gcaatatccc ttcaaacctt gcatgagagt 1140
agaagtagtt gacaagagcg atttatgtcg aacaagagta gcagtgggtg aaagtgaat 1200
tggaggacga ctacggctgg tgtatgaaga gagtgaagat ggaacagacg acttctggtg 1260
ccacatgcac agccccttaa tccaccatat tggatggcca agaagcatag gccatcgatt 1320
caagagatca gatattacga agaaacagga cggacatttc gatacacctc cacacttatt 1380
tgctaaggta aaagaagtag accagagtgg agaatggttc aaagaaggaa tgaatttgga 1440
agctatagac ccattaaatc ttctacaat atgtgttgcc accattagaa aggtgctggc 1500
tgaatgattc ctgatgattg ggattgatgg ctcaagaagc gcagatggat ctgactgggt 1560
ctgttatcat gcaacctctc ctccatttt cccgtgggtt ttctgtgaaa ttaacatgat 1620
agaactgact ccaccagag gtacacaaaa acttcccttt aaatgggttg actacctcag 1680
ggaaaccggc tccattgcag caccagtaaa actatttaat aaggatgttc caaaccacgg 1740
attccggtta ggaatgaat tagaagctgt agatctcatg gagccacggg taatatgtgt 1800
agccacagtt actcgaatta ttcaccatct cttgaggata cattttgatg gttgggaaga 1860
agagtatgac cagtggttag actgtgagtc cctgacctc tatcctgtag ggtggtgtca 1920
gttaactgga tatcaactac agcctccagc atcacagtca tcaagagaaa gccaatcagc 1980
ttcttcaaaa cagaagaaaa aggctaagtc tcagcaatac aaaggacata agaaaatgac 2040
cacgtcacag ctgaaggagg agctgctgga cggggaggac tatagcttcc tccatggagc 2100
atctgaccag gaaagcaatg gctctgccac cgtctacatc aaacaagagc catgaggcga 2160
ctcgaaaacc atgggcaggc ggggctgttt acaggactga tttggaatca gccagctgta 2220
tagcgggcta ttctactggg acattttgct aaacacagaa aaaagttcag ttccagattt 2280
ttcaggtggg ggggaaacta ttttggtggg ggggcaattt tcaattttat aaagacggac 2340
aatttttggt ttgtatttga agcttttgaa agaattttgt aatattttcc aagtttggat 2400
ttatgtgcat tgttaacaag aactgaaatt ataactttt- tggtaagata aaagtttagg 2460
tagcaggatt gaaggaaatg attaagaagg atatagttgt aaatgcacat gaactgtcat 2520
tacaatgaa ccttcttgg- acctgttggg agattttttg gatttttagag ttaggccagt 2580
cacatctcca gcttcccttg ctgcgaaaat atgcaactga acccctcaca gagggctcat 2640
caccatgtag atcacggaag gggtcattaa ttgtgctcgc tgacgtttta ttgcagccca 2700
tttaactggt tgcacagaaa ctttttcatt ctgctaaaat ttattttgag ttgtatatga 2760
aactaggaga actctggata cttttatatg ctctccttca cgtaaagatg attaaaattg 2820
tctaacta- aagtgttaaa ctggaatgg- tgacaaagta cacaagatct cagtctacat 2880
aaagggttgg gggaaattag tattttccta agttttattc gttttccttg ctagaaaaat 2940
cccataaaag ggatattctt gtcagacctc tgccatttcc- tccatctgtg agagagcaac 3000
gtgatgtaac cacacctgag agcagggtact gtcctgtggt acaaaaagaat cacatggaaa 3060
attgttaacc aaaataagtg tagatttcag aatcattcca gctctactct taactgtcct 3120
gaatttgta gaaactgatt tgaagaaat gtttcttaaa attccataca cacacacaca 3180
catattatga attgtcttta attgtagtca gaatttatta taacaatatt tgatttggac 3240
cattttaaac attccctatt ttaaaattca tacggcttcc ttaagaagta gaatgaaagg 3300
gtaaattggt gacatgtttg ctctctgcat tttctaacct tctaccgaat tgtgactgac 3360
tcagagagct ctagcattta ccagtggagt tcagaaacta aactctcagg aattcattgc 3420
atctgttgta gaagtgcctc tgggtctgag cctggcctcc tcagagtggg aactactgcc 3480
acttctctg- gaaaatagcg agggctaata agaaaacta- caaatgatta acctctgcgg 3540
ccctcagcc tttggaatgc taaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3590

```

<210> 56

<211> 631

<212> PRT

<213> Mus musculus

<400> 56

Met Phe Asp Gly Tyr Asp Ser Cys Ser Glu Asp Thr Ser Ser Ser Ser

Page 31 of 56

Arg Phe Lys Arg Ser Asp Ile Thr Lys Lys Gln Asp Gly His Phe Asp
 355 360 365
 Thr Pro Pro His Leu Phe Ala Lys Val Lys Glu Val Asp Gln Ser Gly
 370 375 380
 Glu Trp Phe Lys Glu Gly Met Lys Leu Glu Ala Ile Asp Pro Leu Asn
 385 390 395 400
 Leu Ser Thr Ile Cys Val Ala Thr Ile Arg Lys Val Leu Ala Asp Gly
 405 410 415
 Phe Leu Met Ile Gly Ile Asp Gly Ser Glu Ala Ala Asp Gly Ser Asp
 420 425 430
 Trp Phe Cys Tyr His Ala Thr Ser Pro Ser Ile Phe Pro Val Gly Phe
 435 440 445
 Cys Glu Ile Asn Met Ile Glu Leu Thr Pro Pro Arg Gly Tyr Thr Lys
 450 455 460
 Leu Pro Phe Lys Trp Phe Asp Tyr Leu Arg Glu Thr Gly Ser Ile Ala
 465 470 475 480
 Ala Pro Val Lys Leu Phe Asn Lys Asp Val Pro Asn His Gly Phe Arg
 485 490 495
 Val Gly Met Lys Leu Glu Ala Val Asp Leu Met Glu Pro Arg Leu Ile
 500 505 510
 Cys Val Ala Thr Val Thr Arg Ile Ile His Arg Leu Leu Arg Ile His
 515 520 525
 Phe Asp Gly Trp Glu Glu Glu Tyr Asp Gln Trp Val Asp Cys Glu Ser
 530 535 540
 Pro Asp Leu Asn Pro Val Gly Trp Cys Gln Leu Thr Gly Tyr Gln Leu
 545 550 555 560
 Gln Pro Pro Ala Ser Gln Ser Ser Arg Glu Ser Gln Ser Ala Ser Ser
 565 570 575
 Lys Gln Lys Lys Lys Ala Lys Ser Gln Gln Tyr Lys Gly His Lys Lys
 580 585 590
 Met Thr Thr Ser Gln Leu Lys Glu Glu Leu Leu Asp Gly Glu Asp Tyr
 595 600 605
 Ser Phe Leu His Gly Ala Ser Asp Gln Glu Ser Asn Gly Ser Ala Thr
 610 615 620
 Val Tyr Ile Lys Gln Glu Pro
 625 630

<210> 57
 <211> 1650
 <212> DNA
 <213> Mus musculus

<400> 57
 ggggcttcac tctccagcct tggagagcag tcaggggaagg cctccaggag gaagcagcta 60
 cctgtggcag ggcagagcag tgcctgtacc agtctcccgg aacttcagaa ctgcggtggc 120

```

catggagtct gtggagc cgggtcaaaga cggcatcctc taccagcagc ac aaagtt 180
tggcaagaaa tgctggcgca aagtgtgggc tctgtgtat gcgggaggcc catcagggtg 240
agctcggcta gaaagctggg acgtgctga tgggtggcctg ggaccagcag gcgacaggtc 300
cacaggggcc agccgtcgag gggaacgccg ggtcatacgc ttggctgact gtgtatctgt 360
cctgcctgctg gatggcgaga gctgtcccag ggacactggt gccttcctga ttaccaccac 420
tgagcgaaag cactgttgg ctgcacagca ccgccagtcc tgggtggacc ccatctgtca 480
gctggccttc ccgggtaccg gagaatgttc gtcaggatca ggacaggctg agaatccaaa 540
aaggggcttt gttcccatgg aagaaaactc tatctactcc tcctggcagg aagtgaccga 600
atttcgggtg atcgtgcaga agacagaagc cacctcccgc tgccagctga aaggacccta 660
cctcctggtg ctgggccaaag atgacatcca actgaggag acatccaagc cccaggcctg 720
ttttagctgg ccctaccgtt tcctgcgcaa gtacggctct gacaagggtg tgttctcgtt 780
tgaggttgct cgccgctgtg actcaggtga gggccttttt gccttcagta gcccgctgc 840
cccagacata tgtggggttg tggctgccgc cattgccgc cagcgggagc gtcttcaga 900
gctggccatg tccccacct gccccctgcc tcggggcctc tcctgccct ccttagagcc 960
ccctggagag cttcgggagg tggcccgagg atttgagctg cccactccca gaaagctgac 1020
tctaactgat ccggggcctc aaagcctacc attgtgctc agccccaca caagaaggac 1080
cggcatccgg tctctatgct tccgtgtgca agcagaccag caagcacaca ggcacggcgg 1140
agcatttcta tgagaacgtg tgcatgtgg aggccagcct tgggctgacc aatgggggtc 1200
ctgaagccca agagggcccc cttggtggcc gcagccccct tgggcagcct tatctaccat 1260
aacactgagg atctgagttg gccgggctcg gcccaggaca gcaatctgga agcccagtac 1320
cggaggctgc tggaaactgga gctggatgag gccggaagcg ccggccgctc tggagcgag 1380
gcaggcatca aggccaaagt ggtgaccctg ctgaccctg aacggaagaa gggccccgc 1440
ccctgtgacc ggccctgaag gcctgagcgg ccagccactg caggacagag gtgatcccc 1500
aagaccagga acaacttoga acataaccgg tctactctga cctgcaggga caagccaggt 1560
ggcccgggga ggagccacac tctgccctac ctctccctc agactgtaca gattgaacag 1620
taataaagct tgccatcaa cttcaaaaaa 1650

```

<210> 58

<211> 359

<212> PRT

<213> Mus musculus

<400> 58

```

Met Glu Ser Val Glu Pro Pro Val Lys Asp Gly Ile Leu Tyr Gln Gln
  1                      5                      10                      15

```

```

His Val Lys Phe Gly Lys Lys Cys Trp Arg Lys Val Trp Ala Leu Leu
      20                      25                      30

```

```

Tyr Ala Gly Gly Pro Ser Gly Val Ala Arg Leu Glu Ser Trp Asp Val
    35                      40                      45

```

```

Arg Asp Gly Gly Leu Gly Pro Ala Gly Asp Arg Ser Thr Gly Pro Ser
    50                      55                      60

```

```

Arg Arg Gly Glu Arg Arg Val Ile Arg Leu Ala Asp Cys Val Ser Val
    65                      70                      75                      80

```

```

Leu Pro Ala Asp Gly Glu Ser Cys Pro Arg Asp Thr Gly Ala Phe Leu
      85                      90                      95

```

```

Ile Thr Thr Thr Glu Arg Ser His Leu Leu Ala Ala Gln His Arg Gln
    100                      105                      110

```

```

Ser Trp Val Asp Pro Ile Cys Gln Leu Ala Phe Pro Gly Thr Gly Glu
    115                      120                      125

```

```

Cys Ser Ser Gly Ser Gly Gln Ala Glu Asn Pro Lys Arg Gly Phe Val
    130                      135                      140

```

```

Pro Met Glu Glu Asn Ser Ile Tyr Ser Ser Trp Gln Glu Val Thr Glu
    145                      150                      155                      160

```

```

Phe Pro Val Ile Val Gln Lys Thr Glu Ala Thr Ser Arg Cys Gln Leu

```

161 170 1
 Lys Gly Pro Tyr Leu Leu Val Leu Gly Gln Asp Asp Ile Gln Leu Arg
 180 185 190
 Glu Thr Ser Lys Pro Gln Ala Cys Phe Ser Trp Pro Tyr Arg Phe Leu
 195 200 205
 Arg Lys Tyr Gly Ser Asp Lys Gly Val Phe Ser Phe Glu Ala Gly Arg
 210 215 220
 Arg Cys Asp Ser Gly Glu Gly Leu Phe Ala Phe Ser Ser Pro Arg Ala
 225 230 235 240
 Pro Asp Ile Cys Gly Val Val Ala Ala Ala Ile Ala Arg Gln Arg Glu
 245 250 255
 Arg Leu Pro Glu Leu Ala Met Ser Pro Pro Cys Pro Leu Pro Arg Ala
 260 265 270
 Leu Ser Leu Pro Ser Leu Glu Pro Pro Gly Glu Leu Arg Glu Val Ala
 275 280 285
 Pro Gly Phe Glu Leu Pro Thr Pro Arg Lys Leu Pro Leu Thr Asp Pro
 290 295 300
 Gly Pro Gln Ser Leu Pro Leu Leu Leu Ser Pro Asn Thr Arg Arg Thr
 305 310 315 320
 Gly Ile Arg Ser Leu Cys Val Arg Val Gln Ala Asp Gln Gln Ala His
 325 330 335
 Arg His Gly Gly Ala Phe Leu Glu Arg Val His Ala Gly Gly Gln Pro
 340 345 350
 Trp Ala Asp Gln Trp Gly Ser
 355

<210> 59

<211> 1750

<212> DNA

<213> Mus musculus

<400> 59

ggccctccgac atattgcccg caggagctgc ggcggtgagc ggagagcgcc gggggaagga 60
 gatggggagga cgaagaggtc ccaacaggac atcttactat cgaaaccgcg tctgtgagcc 120
 aggatcatca ggggcctctg gtggaggcca ctcttctagt gcatccgtga gcagcgctcg 180
 ttctcgaagc aggaccactt ctgggacggg cctctccagc cctccgctgg ccgcgcagac 240
 cgttgtgcct ctacagcact gcaagatccc cgagctgccc gtccaggcca gcattctgtt 300
 cgagctacag ctcttcttct gtcagctcat agccctgttc gtgcactaca tcaacatcta 360
 taagacagtc tgggtggtatc caccctcgca cccaccctcc cactctccc tgaacttcca 420
 cctgatcgac ttcaacttgc tgatgggtgac cgccattgtt ctgggcccgc gattcatcgg 480
 gtccatcggt aaggaggctt ctacagagggt gaaagggtct cctcttccgc tccatcctgc 540
 tgttccctac ccgcttcacg ttctcacggc gacaggctgg agtctgtgcc ggtccctcat 600
 ccactcttctc aggacctact ccttccctgaa cctctgttct ctctgtctatc cgtttgggat 660
 gtacattccg ttccctgcagc tgaactatga tcttcgcaag acgaacctct tcacccacat 720
 ggcttccatg ggaccccgag aggcagtcag tggcctggca aggagtcggg actactttct 780
 gacactgcgg gagacttgga agcagcatac acgacagcta tatggcccgc aagccatgcc 840
 caccatgcc tgctgcttgt caccagcct cattcgcaat gaagttgagt tctcctcaaat 900
 ggacttcaat tggcgaatga aggaagtact tgtcagctcc atgctgagtg cctactatgt 960
 ggcctttgta cctgtgtggt ttgtgaagaa tacacattac tatgacaagc gctggtctctg 1020
 tgagctcttc ctgctgtgtg ccatcagtac ctccgtgata ctcatgcagc actgtgtgcc 1080
 tgccagctac tgtgacctgc tgcacaaggc cgccgcccac ctggcggtgct ggcagaaggt 1140

```

ggaccagct ctgtgct atgtgctcca gccaccgtgg actgaagagt gc gtggcc 1200
acaggggtgtg ctggtgaaac acagcaagaa tgtctacaaa gcagtgggcc actacaacgt 1260
ggcatccctt ccgatgtctt ccacttctgc ttccatttct ttttcagcaa cccctgcgg 1320
atcctcaaca tcttttctgt gctggagggg gccgtcattg tctaccagct gtactcctta 1380
atgtcctcag agaagtggca ccagaccatc tccctggccc tcatcctctt cagcaactac 1440
tatgccttct tcaagctgct ccgggaccgc ctggtattgg gcaaagctac tcctactcag 1500
ccagccccca gaggacctgg accaccggtt ctcttgagtc ccaggggat cctccaaaca 1560
acatccgggc ttcaccaggg ggcttctctg ccaagggctc ttgggggttg gggggagggg 1620
tacaacaaac aaaaacaaaa aacaacaaaa aaaaaccaac ccaccagagc ttgtattttt 1680
tgttacgtac tgtttctttt ttgataatt gatgtggtaa gaaaaaagtc ttatttttat 1740
actccaaaaa 1750

```

<210> 60
 <211> 439
 <212> PRT
 <213> Mus musculus

<400> 60
 Met Gly Gly Arg Arg Gly Pro Asn Arg Thr Ser Tyr Tyr Arg Asn Pro
 1 5 10 15
 Leu Cys Glu Pro Gly Ser Ser Gly Ala Ser Gly Gly Gly His Ser Ser
 20 25 30
 Ser Ala Ser Val Ser Ser Val Arg Ser Arg Ser Arg Thr Thr Ser Gly
 35 40 45
 Thr Gly Leu Ser Ser Pro Pro Leu Ala Ala Gln Thr Val Val Pro Leu
 50 55 60
 Gln His Cys Lys Ile Pro Glu Leu Pro Val Gln Ala Ser Ile Leu Phe
 65 70 75 80
 Glu Leu Gln Leu Phe Phe Cys Gln Leu Ile Ala Leu Phe Val His Tyr
 85 90 95
 Ile Asn Ile Tyr Lys Thr Val Trp Trp Tyr Pro Pro Ser His Pro Pro
 100 105 110
 Ser His Thr Ser Leu Asn Phe His Leu Ile Asp Phe Asn Leu Leu Met
 115 120 125
 Val Thr Ala Ile Val Leu Gly Arg Arg Phe Ile Gly Ser Ile Val Lys
 130 135 140
 Glu Ala Ser Gln Arg Gly Lys Gly Leu Pro Leu Pro Leu His Pro Ala
 145 150 155 160
 Val Pro His Pro Leu His Val Leu Thr Ala Thr Gly Trp Ser Leu Cys
 165 170 175
 Arg Ser Leu Ile His Leu Phe Arg Thr Tyr Ser Phe Leu Asn Leu Leu
 180 185 190
 Phe Leu Cys Tyr Pro Phe Gly Met Tyr Ile Pro Phe Leu Gln Leu Asn
 195 200 205
 Tyr Asp Leu Arg Lys Thr Asn Leu Phe Thr His Met Ala Ser Met Gly
 210 215 220
 Pro Arg Glu Ala Val Ser Gly Leu Ala Arg Ser Arg Asp Tyr Phe Leu
 225 230 235 240
 Thr Leu Arg Glu Thr Trp Lys Gln His Thr Arg Gln Leu Tyr Gly Pro

24: 250 4

Glu Ala Met Pro Thr His Ala Cys Cys Leu Ser Pro Ser Leu Ile Arg
260 265 270

Asn Glu Val Glu Phe Leu Lys Met Asp Phe Asn Trp Arg Met Lys Glu
275 280 285

Val Leu Val Ser Ser Met Leu Ser Ala Tyr Tyr Val Ala Phe Val Pro
290 295 300

Val Trp Phe Val Lys Asn Thr His Tyr Tyr Asp Lys Arg Trp Ser Cys
305 310 315 320

Glu Leu Phe Leu Leu Val Ser Ile Ser Thr Ser Val Ile Leu Met Gln
325 330 335

His Leu Leu Pro Ala Ser Tyr Cys Asp Leu Leu His Lys Ala Ala Ala
340 345 350

His Leu Ala Cys Trp Gln Lys Val Asp Pro Ala Leu Cys Ser Asn Val
355 360 365

Leu Gln Pro Pro Trp Thr Glu Glu Cys Met Trp Pro Gln Gly Val Leu
370 375 380

Val Lys His Ser Lys Asn Val Tyr Lys Ala Val Gly His Tyr Asn Val
385 390 395 400

Ala Ser Pro Pro Met Ser Pro Thr Ser Ala Ser Ile Ser Phe Ser Ala
405 410 415

Thr Pro Cys Gly Ser Ser Thr Ser Phe Cys Cys Trp Arg Val Pro Ser
420 425 430

Leu Ser Thr Ser Cys Thr Pro
435

<210> 61
<211> 3071
<212> DNA
<213> Mus musculus

<400> 61
gtgaaagcag cagtgcgcct ctgctccctt cagagcacag cctggtgtca aggtccaggt 60
tccaccggct gctgctgtca ccgcaggga gtctagcccc tcccagaagg agacacagaa 120
gaatggccat ctcaactggg ttgttcctgc tgcctggggct ccttggccag ccctgggcag 180
gggctgtgc tgattcacag gctgtggtgt gcgaggggac tgcctgctat acagccatt 240
ggggcaagct gagtgcgcct gaagcccagc atcgtgcaa tgagaatgga ggcaatctt 300
ccaccgtgaa gagtggagg gagggccggc atgttcagca agccctgact cagctcctga 360
agaccaaggc acccttgaa gcaaagatgg gcaaattctg gatcgggctc cagcgagaga 420
agggcaactg tacgtacat gatttgcaa tgaggggctt cagctgggtg ggtgggtggag 480
aggacacagc ttattcaaac tggtaaaaag ccagcaagag ctctgtatc tttaaactgt 540
gtgtgtccct catactggac ctgtccttga cacctcacc cagcatctg cccaagtggc 600
atgagagtc ctgtgggacc ccgaagctc caggtaacag cattgaagg ttctgtgca 660
agttcaactt caaaggcatg ttagggccac tggcgctggg tggccagggt cgggtgacct 720
ataccacccc ttccaggcc actacctct ctctggaggc tgtgcctttt gctctgtag 780
ccaatgtagc ttgtgggat gaagctaaga gtgaaaccca ctatttccta tgcaatgaaa 840
agactccagg aatatttcac tggggcagct cagggccact ctgtgtcagc cccaagtgtg 900
gttgacagtt caacaacggg ggctgccagc aggtattgctt cgaaggtggc gatggctcct 960
tccgtgcgg ctgcccggct ggatttcgac tgcctggatga tctagtaact tgtgcttcca 1020
ggaaacctcg cagctcaaac ccatgcacag gaggtggcat gtgccattct gtaccactca 1080
gtgaaaacta cacttgccgt tgtcccagcg gctaccagct ggactctagc caagtgcact 1140

```

gtgtggatat agatgag caggactccc cctgtgcccc ggattgtgtc aa ctctag 1200
ggagcttcca ctgtgaatgt tgggttgggt accaaccacag tggccccaag gaagaggcct 1260
gtgaagatgt ggaatgagtg gcagctgcca actcgccctg tgcccaaggc tgcatcaaca 1320
ctgatggctc tctctactgc tcctgtaaag agggctatat tgtgtctggg gaagacagta 1380
ccagtggtga ggaatagat gagtgttcgg acgcaagggg caatccatgt gattccctgt 1440
gcttcaacac agatgggtcc ttcaggtgtg gctgcccggc aggctgggag ctggctccca 1500
atggggctct tctagcagg ggcactgtgt tttctgaact accagccagg cctcccaaaa 1560
aggaagacaa cgatgacaga aaggagagta ctatgcctcc tactgaaatg cccagtcttc 1620
ctagtggctc taaggatgtc tccaacagag cacagacaac aggtctcttc gtccaatcag 1680
atattcccac tgcctctgtt ccactagaaa tagaaatccc tagtgaagta tctgatgtct 1740
ggttcgagtt gggcacatac ctccccacga cctccggcca cagcaagccg acacatgaag 1800
attctgtgtc tgcacacagt gacaccgatg ggcagaacct gcttctgttt tacatcctgg 1860
ggacgggtgg gccatctca ctcttgctgg tgcctggcct agggattctc atttatcata 1920
aacggagagc caagaaggag gagataaaag agaagaagcc tcagaatgca gccgacagct 1980
attcttgggt tcagagcgga gcagagagcc aagccccgga gaatcagtac agccaacac 2040
cagggacaga ctgctgaaga ctatgtggcc ttagagacag ctgccactac cttcagagct 2100
accttcttag atgaggggga agccacatca ttctgaatga cttgactgga ctctcagcaa 2160
aaaaattgtg cacttccac ttaagaacct ggtggcttgg gataggcagg tattttcttg 2220
gtgcctttga tatgtctggg ggtgaaagct gtgtgttggg ttgtcattgt ggggagtttt 2280
gtggatattg acagacctca ctcaaacacc cttttcaaat ccaatagcaa ctggttcctc 2340
tgggtcctaa ttagggggaa aggagtcaga ggggtgggac aggggtgggg gatggggctt 2400
caaaagtttt tcttatcact tgatttatca tcgaaggagt tactgtgtct aattacaatg 2460
gaaacagttc ctttccatca caggacagac acacctcaat cctccatggg gtcaacaaat 2520
atataccccc agtgacccct taggcaagga cttgttgaga actgcacac attttgacct 2580
gttctcaaca gtacccatct atttcaggtg ggatctcttg acctttctc cttcccatct 2640
tgtctgcaat gtggcaaatg gcttctttt gcatttttac tccgccccca cccaagctg 2700
aagttcattt gcagatcagc gattaagtct gaattgtgtg gtggtcagtc ttgtttcctt 2760
ttgtcagggg ttattgtaaa tgttagtaat ttgcctcaa gccctcagta agaacataaa 2820
tattttaaaa tatgtgcgtt tgaaatctgt ttcatgcac ctggaactgt gggatgtctc 2880
ggcaagagtg actttagtct ttcagtgaat gttgcccaga atgtgggtag ggaaggctca 2940
caggttactc tcttcttag agctacaaca taacattctg aggggagtca cagggttgcc 3000
tttaaaaagt gggagctatg tcatgctttg agctttctgt taagcacctc tcctaataaa 3060
ctctgaaaaa a 3071

```

<210> 62
 <211> 644
 <212> PRT
 <213> Mus musculus

<400> 62
 Met Ala Ile Ser Thr Gly Leu Phe Leu Leu Leu Gly Leu Leu Gly Gln
 1 5 10 15
 Pro Trp Ala Gly Ala Ala Ala Asp Ser Gln Ala Val Val Cys Glu Gly
 20 25 30
 Thr Ala Cys Tyr Thr Ala His Trp Gly Lys Leu Ser Ala Ala Glu Ala
 35 40 45
 Gln His Arg Cys Asn Glu Asn Gly Gly Asn Leu Ala Thr Val Lys Ser
 50 55 60
 Glu Glu Glu Ala Arg His Val Gln Gln Ala Leu Thr Gln Leu Leu Lys
 65 70 75 80
 Thr Lys Ala Pro Leu Glu Ala Lys Met Gly Lys Phe Trp Ile Gly Leu
 85 90 95
 Gln Arg Glu Lys Gly Asn Cys Thr Tyr His Asp Leu Pro Met Arg Gly
 100 105 110
 Phe Ser Trp Val Gly Gly Gly Glu Asp Thr Ala Tyr Ser Asn Trp Tyr
 115 120 125

Lys Ala Ser Lys Ser ar Cys Ile Phe Lys Arg Cys Val Ser i Ile
 130 135 140
 Leu Asp Leu Ser Leu Thr Pro His Pro Ser His Leu Pro Lys Trp His
 145 150 155 160
 Glu Ser Pro Cys Gly Thr Pro Glu Ala Pro Gly Asn Ser Ile Glu Gly
 165 170 175
 Phe Leu Cys Lys Phe Asn Phe Lys Gly Met Cys Arg Pro Leu Ala Leu
 180 185 190
 Gly Gly Pro Gly Arg Val Thr Tyr Thr Thr Pro Phe Gln Ala Thr Thr
 195 200 205
 Ser Ser Leu Glu Ala Val Pro Phe Ala Ser Val Ala Asn Val Ala Cys
 210 215 220
 Gly Asp Glu Ala Lys Ser Glu Thr His Tyr Phe Leu Cys Asn Glu Lys
 225 230 235 240
 Thr Pro Gly Ile Phe His Trp Gly Ser Ser Gly Pro Leu Cys Val Ser
 245 250 255
 Pro Lys Phe Gly Cys Ser Phe Asn Asn Gly Gly Cys Gln Gln Asp Cys
 260 265 270
 Phe Glu Gly Gly Asp Gly Ser Phe Arg Cys Gly Cys Arg Pro Gly Phe
 275 280 285
 Arg Leu Leu Asp Asp Leu Val Thr Cys Ala Ser Arg Asn Pro Cys Ser
 290 295 300
 Ser Asn Pro Cys Thr Gly Gly Gly Met Cys His Ser Val Pro Leu Ser
 305 310 315 320
 Glu Asn Tyr Thr Cys Arg Cys Pro Ser Gly Tyr Gln Leu Asp Ser Ser
 325 330 335
 Gln Val His Cys Val Asp Ile Asp Glu Cys Gln Asp Ser Pro Cys Ala
 340 345 350
 Gln Asp Cys Val Asn Thr Leu Gly Ser Phe His Cys Glu Cys Trp Val
 355 360 365
 Gly Tyr Gln Pro Ser Gly Pro Lys Glu Glu Ala Cys Glu Asp Val Asp
 370 375 380
 Glu Cys Ala Ala Ala Asn Ser Pro Cys Ala Gln Gly Cys Ile Asn Thr
 385 390 395 400
 Asp Gly Ser Phe Tyr Cys Ser Cys Lys Glu Gly Tyr Ile Val Ser Gly
 405 410 415
 Glu Asp Ser Thr Gln Cys Glu Asp Ile Asp Glu Cys Ser Asp Ala Arg
 420 425 430
 Gly Asn Pro Cys Asp Ser Leu Cys Phe Asn Thr Asp Gly Ser Phe Arg
 435 440 445
 Cys Gly Cys Pro Pro Gly Trp Glu Leu Ala Pro Asn Gly Val Phe Cys
 450 455 460
 Ser Arg Gly Thr Val Phe Ser Glu Leu Pro Ala Arg Pro Pro Gln Lys

```
<210> 63
<211> 499
<212> DNA
<213> Mus musculus
```

[illegible]

```
<210> 64
<211> 89
<212> PRT
<213> Mus musculus
```

<400> 64
Met Arg Thr Ala Pro Ala Ala Ala Ile Leu Cys Glu Leu Gly Leu Leu
1 5 10 15
Tyr Gly Asn Gly Val Arg Lys Cys Gly Asn His Lys Arg Gln
20 25 30

Asn Gln Lys Thr Glu Gly Ser Trp Ala Met Ser His Ser Lys Glu Ser
 35 40 45
 Leu Pro Pro Asn Ile Cys Thr Gly Leu Leu Gln Glu Gln Glu Met Asn
 50 55 60
 Tyr Arg Val Gln Pro Leu Arg Leu Gly Leu Asn Met Ile Phe Thr Leu
 65 70 75 80
 Ser Thr Leu Pro Asn Arg Ala Glu Met
 85

<210> 65
 <211> 3857
 <212> DNA
 <213> Mus musculus

<400> 65
 agagggtaga cagaaaggcg ggaagggt gtgaggtcaa atggaccca tggaaactgaa 60
 caacgtcagc atcgaaccgc acggagacag ctgcagcggg gacagtattc aggacagcta 120
 caccggcatg gaaaactccg acaaggacgc catgaacagc caatttgcta atgaagatgc 180
 cgaaagtcag aagttctctga caaatgggtt tttaggggaag aagaagctag ccgattacgc 240
 ggatgagcat caccctggaa tgacttcctt tggaaatgtcc tcatttaacc tgagcaacgc 300
 catcatgggc agtgggatct taggcttgct ctatgccatg gccaacaccg ggatcatcct 360
 ttttataatc atgctgctta ctgtggcaat actctcgctc tactcggttc accctttgct 420
 gaagacagcc aagggaaggag ggtctctaata ctatgaaaaa ttgggcgaga aagcatttgg 480
 atggcctggg aaaattggag ccttcattctc tattacaatg cagaacattg gagccatgtc 540
 aagctacctc ttcattcatta agtacgaact gcctgaagta atcagagcat tcatgggact 600
 tgaagaaaac actggggaat ggtacctcaa cggcaactac ctctcttat ttgtgtccgt 660
 ggggatcatc ctcccgtctc ctctccttaa aaatttaggc taccttggct acaccagtgg 720
 attttctctc tcttgcatgg tgttttctgt cagtgtggtg atttataaaa aattccaaat 780
 tccctgccct ctgctctctc tggatcacaa caacggaaat ctgacgttca acaacacact 840
 tccgattcac atgatctcgc tgcctaataga ctcgagagc tcgggtgtga acttcatgat 900
 ggattacgct caccacaacc cagctgggct ggatgagaag caggtcgcag gccctcttca 960
 cagcaatggc gtggagtagc aagcccgagg tgcagagaaa tgccaaccaa aatactttgt 1020
 gttcaattcc cggacggccc atgcaatccc aatcctggct tttgcttttg tctgccacc 1080
 tgaggtcctt cccatctaca gcgagcttaa agatcgatcc cgcagaaaga tgcagacgg 1140
 gtccaacatt tccatctcag gcaatgctct catgtacctt ctgctggccc tcttgggtta 1200
 tctgagcttc tacggggacg ttgaagacga gctgctgcat gcttacagca aggtctacac 1260
 atttgatacg gctctcttca tgggtcgccct ggagctcctg gtggcagtga cactgaccgt 1320
 gcccacgtg ctgttctccga tccgtacttc ggtgatcaca ctgctgtttc caaggaaacc 1380
 cttcagctgg ctgaagcatc tggggatcgc tgcaatcatc atcgactca acaacatct 1440
 ggtcatcctc gtgcctacca tcaaatacat ctttggtatc ataggggctt cttctgccac 1500
 tatgtctgatt ttcattcttc cggctgcgtt ttatctcaag ctctgcaaga aagaacctct 1560
 aagatcaccc cagaagattg gggctttggt ctctcttggt actggaatta ttttcatgat 1620
 gggaaagcatg gcgtctcata tactcgactg gatctacaac ccgccgaatc ccaatcacca 1680
 ctaatcccgc ggagacgcgt ctccactgga aacagctgaa attgtctgaa ggacatttta 1740
 gttgtcttga ttgggaatgt agtctgagga attagcaaga ttccaaagac gtttttctag 1800
 ctctatcagc acacatttta acccaggccg tgcagtgcag tgtgtgatgc ccgagttgtg 1860
 tttgcagcag ctgtgcaagc tgaagcctgt tggctgcgtg tgttggctag cagacaatag 1920
 cctgtccccc catggtaact ccactctctc ccacccccag attaacaggt aattctactc 1980
 tcagaacatc agacaaaagac ctctcgtgtg ggatacttgt ggaagagaaa attatgggtt 2040
 ttgttgggaa tgggttttgtt gggaatggtg aaggatgcat taaaaattct gtgcgaagta 2100
 tcatcagtta cggccatctc tcaactctaca ccaacactaa gggctcgggtt actagctgag 2160
 gcagggggat atcttgggct gtcctctgtg ggatcatgac gtatgacggg tgccagtata 2220
 gagtacttca tttcaataact caaggaatag tttgcccac ctgcttatta caccgagtta 2280
 gtgaatttct attcttttta ttttttggta tggcaaatgt tcaaatactc cactaaacat 2340
 agaggggggt tattttagttt atcatgaaac aactaccaaa aagtaatggt ttttaacatc 2400
 tgccctttca tgttgtttct gatatctttg ttccactttg tctttgaaca aggtctcccc 2460
 tctcgtcttc cgttctcact ctctcgtaatt ttcagaaaaa ccacagtcgg tgtgggagac 2520
 acactacccc agtattgttt gatatactc tatttgataa acattcagtg caggaaactg 2580
 tgattttgct atatgtttgt gtacataatc tcattctgca gttatcagaa cgttgacata 2640
 tgggacattg gattttttat ttttatacat gtaggttttt tttcttcaca gacaaaaatg 2700

```

ttatatcatc aggggtgggc aggggaattaa gctggtgggc tcaaaaatcc atgggtactt 2760
atctgtccat tggagacatc taaaaattraa agtcaaagtt gtgcatagtt cagtaatgct 2820
cttcactgtt tacaagacta taatcatctc agcaaagtag cgaagggaagg gtttgtaaat 2880
aattttcagt gctaacaatg gtctttgaat ttaacatgtc tagaataggg tttagtccat 2940
tttagtttac atcggaactt ggtgacattc atttgccatt aatagaacat cgtgtgggta 3000
ggaataaatg caccaacatg acccaaaacc acattgctca catccatgga cgcttacacg 3060
ggatggagtt cagactccgc ccactgccta gaatctctgg cattgacttc ttttttcaa 3120
ttactgtgtg tatattttac atttcagtaa aaagccaagt ctgaaatgca agtatccctac 3180
aggtagtgat atactgtcac accaaagcag tttttcaaca gtgtgcctaa gaaacatata 3240
ttgttttact gtgtcatttt ctggacctta agactataca atgggtgtgct ttgtgataga 3300
aattgagcat tactagaaac taccataact tcggagttat tatcatttgg gagggaaaaa 3360
aatcaccaaa taggatcttt aagatgtctt tctattttca ttcaccaag aaataaatca 3420
agatagaag cagaaagggt gtgctgcaaa tattaaatta ttattttatc tgaacatatt 3480
aatgtgaata taaaagatta atccaaaaat attggaaacc tgtgaaatct gtacgtattt 3540
cggatcttta cacttagagc agctaacctg gaagtatcca cttgtgtagt tagatattca 3600
agagcaatat cacagtgcga aagacttcat tcacacagca ccgatcattt gtgtggccgt 3660
tttcaggatt agagttcaca cagactccag ggcagatact cagaataaccg ttttctggta 3720
aatttaaccc atttgtaaac agatacaaat tttattttct tacatcattt ataagacggt 3780
tcaatgtact ggggggattt tttttattac agactgtgta ttggtatata ataaataaac 3840
ttttcaaatg aaaaaaa 3857

```

<210> 66

<211> 547

<212> PRT

<213> Mus musculus

<400> 66

```

Met Asp Pro Met Glu Leu Asn Asn Val Ser Ile Glu Pro Asp Gly Asp
 1           5           10          15

Ser Cys Ser Gly Asp Ser Ile Gln Asp Ser Tyr Thr Gly Met Glu Asn
 20          25          30

Ser Asp Lys Asp Ala Met Asn Ser Gln Phe Ala Asn Glu Asp Ala Glu
 35          40          45

Ser Gln Lys Phe Leu Thr Asn Gly Phe Leu Gly Lys Lys Lys Leu Ala
 50          55          60

Asp Tyr Ala Asp Glu His His Pro Gly Met Thr Ser Phe Gly Met Ser
 65          70          75          80

Ser Phe Asn Leu Ser Asn Ala Ile Met Gly Ser Gly Ile Leu Gly Leu
 85          90          95

Ser Tyr Ala Met Ala Asn Thr Gly Ile Ile Leu Phe Ile Ile Met Leu
100         105         110

Leu Thr Val Ala Ile Leu Ser Leu Tyr Ser Val His Leu Leu Leu Lys
115         120         125

Thr Ala Lys Glu Gly Gly Ser Leu Ile Tyr Glu Lys Leu Gly Glu Lys
130         135         140

Ala Phe Gly Trp Pro Gly Lys Ile Gly Ala Phe Ile Ser Ile Thr Met
145         150         155         160

Gln Asn Ile Gly Ala Met Ser Ser Tyr Leu Phe Ile Ile Lys Tyr Glu
165         170         175

Leu Pro Glu Val Ile Arg Ala Phe Met Gly Leu Glu Glu Asn Thr Gly
180         185         190

Glu Trp Tyr Leu Asn Gly Asn Tyr Leu Val Leu Phe Val Ser Val Gly

```

195	200	205
Ile Ile Leu Pro Leu Ser Leu Leu Lys Asn Leu Gly Tyr Leu Gly Tyr		
210	215	220
Thr Ser Gly Phe Ser Leu Ser Cys Met Val Phe Phe Val Ser Val Val		
225	230	235 240
Ile Tyr Lys Lys Phe Gln Ile Pro Cys Pro Leu Pro Ala Leu Asp His		
	245	250 255
Asn Asn Gly Asn Leu Thr Phe Asn Asn Thr Leu Pro Ile His Met Ile		
	260	265 270
Ser Leu Pro Asn Asp Ser Glu Ser Ser Gly Val Asn Phe Met Met Asp		
	275	280 285
Tyr Ala His His Asn Pro Ala Gly Leu Asp Glu Lys Gln Val Ala Gly		
	290	295 300
Pro Leu His Ser Asn Gly Val Glu Tyr Glu Ala Gln Gly Ala Glu Lys		
305	310	315 320
Cys Gln Pro Lys Tyr Phe Val Phe Asn Ser Arg Thr Ala Tyr Ala Ile		
	325	330 335
Pro Ile Leu Ala Phe Ala Phe Val Cys His Pro Glu Val Leu Pro Ile		
	340	345 350
Tyr Ser Glu Leu Lys Asp Arg Ser Arg Arg Lys Met Gln Thr Val Ser		
	355	360 365
Asn Ile Ser Ile Ser Gly Met Leu Val Met Tyr Leu Leu Ala Ala Leu		
	370	375 380
Phe Gly Tyr Leu Ser Phe Tyr Gly Asp Val Glu Asp Glu Leu Leu His		
385	390	395 400
Ala Tyr Ser Lys Val Tyr Thr Phe Asp Thr Ala Leu Leu Met Val Arg		
	405	410 415
Leu Ala Val Leu Val Ala Val Thr Leu Thr Val Pro Ile Val Leu Phe		
	420	425 430
Pro Ile Arg Thr Ser Val Ile Thr Leu Leu Phe Pro Arg Lys Pro Phe		
	435	440 445
Ser Trp Leu Lys His Phe Gly Ile Ala Ala Ile Ile Ile Ala Leu Asn		
	450	455 460
Asn Ile Leu Val Ile Leu Val Pro Thr Ile Lys Tyr Ile Phe Gly Phe		
465	470	475 480
Ile Gly Ala Ser Ser Ala Thr Met Leu Ile Phe Ile Leu Pro Ala Ala		
	485	490 495
Phe Tyr Leu Lys Leu Val Lys Lys Glu Pro Leu Arg Ser Pro Gln Lys		
	500	505 510
Ile Gly Ala Leu Val Phe Leu Val Thr Gly Ile Ile Phe Met Met Gly		
	515	520 525
Ser Met Ala Leu Ile Ile Leu Asp Trp Ile Tyr Asn Pro Pro Asn Pro		
	530	535 540

Asn His His
545

<210> 67
<211> 1070
<212> DNA
<213> Mus musculus

<400> 67
agcgacatgg ccccgcccg cgtccaggcc cagcctccag gcggctctca actgagggttc 60
ctgctgttcc tgctgtgttt gctgctgctg ctgtcatggc catcgagggg ggacgccctg 120
gcaatgcctg aacagcgacc ctccggccct gaggcccaac tcaacgccga cgagctacgg 180
ggtcgcttcc aggacctgct gagccggctg catgccaaac agagccgaga ggactcgaac 240
tcagaaccaaa gtccctgaccc agctgtccgg ataactcagtc cagagggtgag attgggggtcc 300
cacggccagc tgctactccg cgtcaaccgg gcgtcgctga gtcagggtct ccccgaaagcc 360
taccggtgac accgagcgct gctcctgctg acgcccagcg cccggccctg ggacatcact 420
aggcccttga agcgtgcgct cagcctccgg ggaccccggtg ctcccgcatc acgctcgcg 480
ctgacgccgc ctccggacct ggctatgctg ccctctggcg gcacgcagct ggaactgcgc 540
ttacgggtag ccgcccggcag ggggcgcgga agcgcgcatg cgcacccaag agactcgtgc 600
ccactgggtc cggggcgctg ctgtcacttg gagactgtgc aggcaactct tgaagacttg 660
ggctggagcg actgggtgct gtccccgcgc cagctgcagc tgagcatgtg cgtggggcag 720
tgtccccacc tgtatcgctc cgcgaacacg catgcgcaga tcaaaagcacg cctgcatggc 780
ctgcagcctg acaagggtgc tgccccgtgc tgtgtccctt ccagctacac cccgggtggtt 840
cttatgcaca ggacagacag tgggtgtgtca ctgcagactt atgatgacct ggtggcccg 900
ggctgccact gcgcttgagc accgggcccct gctcctcacc tacactcccc ttcaaggatg 960
ctatttatat ttgtatttat taattatttt aatttatttg ggtcgggctg ggtggatgga 1020
ttgtgtattt atttaaaact ctgctaataa aggtgagctt ggtttcaaaa 1070

<210> 68
<211> 298
<212> PRT
<213> Mus musculus

<400> 68
Met Ala Pro Pro Ala Leu Gln Ala Gln Pro Pro Gly Gly Ser Gln Leu
1 5 10 15
Arg Phe Leu Leu Phe Leu Leu Leu Leu Leu Leu Leu Ser Trp Pro
20 25 30
Ser Gln Gly Asp Ala Leu Ala Met Pro Glu Gln Arg Pro Ser Gly Pro
35 40 45
Glu Ser Gln Leu Asn Ala Asp Glu Leu Arg Gly Arg Phe Gln Asp Leu
50 55 60
Leu Ser Arg Leu His Ala Asn Gln Ser Arg Glu Asp Ser Asn Ser Glu
65 70 75 80
Pro Ser Pro Asp Pro Ala Val Arg Ile Leu Ser Pro Glu Val Arg Leu
85 90 95
Gly Ser His Gly Gln Leu Leu Leu Arg Val Asn Arg Ala Ser Leu Ser
100 105 110
Gln Gly Leu Pro Glu Ala Tyr Arg Val His Arg Ala Leu Leu Leu Leu
115 120 125
Thr Pro Thr Ala Arg Pro Trp Asp Ile Thr Arg Pro Leu Lys Arg Ala
130 135 140
Leu Ser Leu Arg Gly Pro Arg Ala Pro Ala Leu Arg Leu Arg Leu Thr

145 50 155 160
 Pro Pro Pro Asp Leu Ala Met Leu Pro Ser Gly Gly Thr Gln Leu Glu
 165 170 175
 Leu Arg Leu Arg Val Ala Ala Gly Arg Gly Arg Arg Ser Ala His Ala
 180 185 190
 His Pro Arg Asp Ser Cys Pro Leu Gly Pro Gly Arg Cys Cys His Leu
 195 200 205
 Glu Thr Val Gln Ala Thr Leu Glu Asp Leu Gly Trp Ser Asp Trp Val
 210 215 220
 Leu Ser Pro Arg Gln Leu Gln Leu Ser Met Cys Val Gly Glu Cys Pro
 225 230 235 240
 His Leu Tyr Arg Ser Ala Asn Thr His Ala Gln Ile Lys Ala Arg Leu
 245 250 255
 His Gly Leu Gln Pro Asp Lys Val Pro Ala Pro Cys Cys Val Pro Ser
 260 265 270
 Ser Tyr Thr Pro Val Val Leu Met His Arg Thr Asp Ser Gly Val Ser
 275 280 285
 Leu Gln Thr Tyr Asp Asp Leu Val Ala Arg
 290 295

<210> 69
 <211> 134
 <212> PRT
 <213> Mus musculus

<400> 69
 Ala Lys Tyr Val Cys Leu Ala Asn Lys Asn Cys Pro Val Asp Lys Arg
 1 5 10 15
 Arg Arg Asn Arg Cys Gln Tyr Cys Arg Phe Gln Lys Cys Leu Ala Val
 20 25 30
 Gly Met Val Lys Glu Val Val Arg Thr Asp Ser Leu Lys Gly Arg Arg
 35 40 45
 Gly Arg Leu Pro Ser Lys Pro Lys Ser Pro Gln Glu Pro Ser Pro Pro
 50 55 60
 Ser Pro Pro Phe Gln Ala Asn Pro Asp Tyr Gln Met Ser Gly Asp Asp
 65 70 75 80
 Thr Gln His Ile Gln Gln Phe Tyr Asp Leu Leu Thr Gly Ser Met Glu
 85 90 95
 Ile Ile Arg Gly Trp Ala Glu Lys Ile Pro Gly Phe Ala Asp Leu Pro
 100 105 110
 Lys Ala Asp Gln Asp Leu Leu Phe Glu Ser Ala Phe Leu Glu Leu Phe
 115 120 125
 Val Leu Arg Leu Ala Tyr
 130

<210> 70
 <211> 140
 <212> PRT
 <213> Mus musculus

<400> 70
 Ser Gln Thr Arg Gln Gln Gly Pro Leu Arg Ser Ile Met Lys Asp Leu
 1 5 10 15
 His Ser Asp Asp Asn Glu Glu Glu Ser Asp Glu Val Glu Asp Asn Asp
 20 25 30
 Asn Asp Ser Glu Met Glu Arg Pro Val Asn Arg Gly Xaa Ser Arg Ser
 35 40 45
 Arg Arg Val Ser Leu Ser Asp Gly Ser Asp Ser Glu Ser Ser Ser Ala
 50 55 60
 Ser Ser Pro Leu His His Glu Pro Pro Pro Pro Leu Leu Lys Thr Asn
 65 70 75 80
 Asn Asn Gln Ile Leu Glu Val Lys Ser Pro Ile Lys Gln Ser Lys Ser
 85 90 95
 Asp Lys Gln Ile Lys Asn Gly Glu Cys Asp Lys Ala Tyr Leu Asp Glu
 100 105 110
 Leu Val Glu Leu Pro Arg Xaa Xaa Met Thr Leu Arg Glu Arg His Ile
 115 120 125
 Leu Gln Gln Ile Val Asn Leu Xaa Glu Glu Thr Gly
 130 135 140

<210> 71
 <211> 89
 <212> PRT
 <213> Mus musculus

<400> 71
 Met Arg Thr Ala Pro Ala Ala Ala Ile Leu Cys Glu Leu Gly Leu Leu
 1 5 10 15
 Tyr Gly Asn Gly Gly Val Arg Lys Cys Gly Asn His Lys Arg Arg Gln
 20 25 30
 Asn Gln Lys Thr Glu Gly Ser Trp Ala Met Ser His Ser Lys Glu Ser
 35 40 45
 Leu Pro Pro Asn Ile Cys Thr Gly Leu Leu Gln Glu Gln Glu Met Asn
 50 55 60
 Tyr Arg Val Gln Pro Leu Arg Leu Gly Leu Asn Met Ile Phe Thr Leu
 65 70 75 80
 Ser Thr Leu Pro Asn Arg Ala Glu Met
 85

<210> 72
 <211> 3121
 <212> DNA
 <213> Homo sapiens

<400> 72

```

gaaagcagca gtgcgcctct gctcccttca gagcacagcc tgggtgtcaag gtccaggttc 60
caccggctgc tgctgtcacc gcaggggagt ctagccctcc ccagaaggag acacagaaga 120
atggccatct caactgggtt gtccctgctg ctggggctcc ttggccagcc ctgggcaggg 180
gctgctgctg attcacaggc tgtggtgtgc gaggggactg cctgctatac agcccatgtg 240
ggcaagctga gtgcgcctga agcccagcat cgtgcaatg agaattggagg caatcttgcc 300
accgtgaaga gtgaggagga ggcccggcat gtccagcaag ccctgactca gctcctgaag 360
accaaggcac ccttggaagc aaagatgggc aaattctgga tcgggctcca gcgagagaag 420
ggcaactgta cgtaccatga ttgccaatg aggggcttca gctgggtggg tgggtggagag 480
gacacagctt attcaaactg gtacaaagcc agcaagagct cctgtatctt taaacgctgt 540
gtgtccctca tactggacct gtccttgaca cctcacccca gccatctgcc caagtggcat 600
gagagtcctt gtgggacccc cgaagctcca ggtaacagca ttgaagggtt cctgtgcaag 660
ttcaacttca aaggcatgtg tagggcactg gcgctgggtg gtccagggcg ggtgacctat 720
accaccctct tccaggccac tacctcctct ctggaggctg tgccctttgc ctctgtagcc 780
aatgtagctt gtggggatga agctaagagt gaaacccact atttccatg caatgaaaaa 840
actccaggaa tatttcaactg gggaagctca gcccactct gtgtcagccc caagtgttgt 900
tgcaagtttca acaacggggg ctgccagcag gattgcttcg aaggtggcga tggctccttc 960
cgctgcgggt gccggcctgg atttcgactg ctggatgatc tagtaacttg tgcctccagg 1020
aacccttgca gctcaaaccc atgcacagga ggtggcatgt gccattctgt accactcagt 1080
gaaaaactaca cttgccgttg tcccagcgcc taccagctgg actctagcca agtgactgt 1140
gtggatatag atgagtgcga ggactcccc tgtgccagg atttgttcaa cactctaggg 1200
agcttccact gtgaatgttg ggttgggtac caaccagtg gccccaagga agaggcctgt 1260
gaagatgttg atgagtgtgc agctgccaac tcgccctgtg cccaaggctg catcaacact 1320
gatggctctt tctactgttc ctgtaaagag ggctatatg tgtctgggga agacagtacc 1380
cagtgtagag atatatagta gtgttcggac gcaaggggca atccatgtga ttccctgtgc 1440
ttcaacacag atggttcctt caggtgtggc tgcccgcag gctgggagct ggctcccaat 1500
ggggtctttt gtagcagggg cactgtgttt tctgaactac cagccaggcc tccccaaaag 1560
gaagacaacg atgacagaaa ggagagtact atgcctccta ctgaaatgcc cagttctcct 1620
agtggctcta aggatgtctc caacagagca cagacaacag gtctcttcgt ccaatcagat 1680
attcccactg cctctgttcc actagaaata gaaatcccta gtgaagtac tgatgtctgg 1740
ttcagagtgg gcacatacct ccccacgacc tccggccaca gcaagccgac acatgaagat 1800
tctgtgtctg cacacagtga caccgatggg cagaacctgc ttctgtttta catcctgggg 1860
acgggtgggtg ccatctcact cttgctggtg ctggccctag ggatttctat ttatcataaa 1920
cggagagcca agaaggagga gataaaagag aagaagcctc agaatgcagc cgacagctat 1980
tcttgggttc cagagcgagc agagagccaa gccccggaga atcagtacag cccaacacca 2040
gggacagact gctgaagact atgtggcctt agagacagct gccactacct tcagagctac 2100
cttcttagat gagggggaag ccacatcatt ctgaatgact tgactggact ctcagcaaaa 2160
aaattgtgca ccttccactt aagaacctgg tggcttggga taggcaggta ttttcttgtt 2220
gcctttgata tgtctggggg tgaagctgt gtgttgggtt gtcattgtgg ggagttttgt 2280
ggatattgac agacctcact caaacacct tttcaaatcc aatagcaact ggttcctctg 2340
gttccctaatt agggggaaaag gagtcaagg ggtgggacag ggtgggggga tggggcttca 2400
aagttttttc ttatcacttg atttatcatc gaaggagtta ctggtgctaa ttacaatgga 2460
aacagttcct ttccatcaca ggacagacac acctcaatcc tccatggggg caacaactat 2520
atacccccag tgacccctta ggcaaggact tgttgagaac tgcatacat tttgacctgt 2580
tctcaacagt acccatctat ttcaggtggg atctctggac ctttccctct tcccatcttg 2640
tctgcaatgt ggcaaatggc ttctttttgc atttttactc cgccccacc ccaagctgaa 2700
gttcatttgc agatcagcga ttaagtctga attgtgtggt ggtcagctct gtttcccttt 2760
gtcagggggt attgtaaatg ttagtaattt cgcctcaagc cctcagraag aacataaata 2820
ttttaaaata tgtgcgtttg aaatctgttt catgcacact ggaactgtgg gatgctcagg 2880
caagagtgc tttagtcttt cagtgaatgt tgcccagaat gtgggtaggg aaggctcaca 2940
ggttactctc ctcccttagag ctacaacata acattctgag gggagtcaca ggggtgcctt 3000
taaaaagtgg gagctatgtc atgctttgag ctttctgtta agcacctctc ctaataaact 3060
ctgaaaaaat aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3120
a

```

<210> 73

<211> 644

<212> PRT

<213> Homo sapiens

<400> 73

```

Met Ala Ile Ser Thr Gly Leu Phe Leu Leu Leu Gly Leu Leu Gly Gln
1           5           10          15

```

Pro Trp Ala Gly Ala Ala Ala Asp Ser Gln Ala Val Val Cys Gly
 20 25 30
 Thr Ala Cys Tyr Thr Ala His Trp Gly Lys Leu Ser Ala Ala Glu Ala
 35 40 45
 Gln His Arg Cys Asn Glu Asn Gly Gly Asn Leu Ala Thr Val Lys Ser
 50 55 60
 Glu Glu Glu Ala Arg His Val Gln Gln Ala Leu Thr Gln Leu Leu Lys
 65 70 75 80
 Thr Lys Ala Pro Leu Glu Ala Lys Met Gly Lys Phe Trp Ile Gly Leu
 85 90 95
 Gln Arg Glu Lys Gly Asn Cys Thr Tyr His Asp Leu Pro Met Arg Gly
 100 105 110
 Phe Ser Trp Val Gly Gly Gly Glu Asp Thr Ala Tyr Ser Asn Trp Tyr
 115 120 125
 Lys Ala Ser Lys Ser Ser Cys Ile Phe Lys Arg Cys Val Ser Leu Ile
 130 135 140
 Leu Asp Leu Ser Leu Thr Pro His Pro Ser His Leu Pro Lys Trp His
 145 150 155 160
 Glu Ser Pro Cys Gly Thr Pro Glu Ala Pro Gly Asn Ser Ile Glu Gly
 165 170 175
 Phe Leu Cys Lys Phe Asn Phe Lys Gly Met Cys Arg Pro Leu Ala Leu
 180 185 190
 Gly Gly Pro Gly Arg Val Thr Tyr Thr Thr Pro Phe Gln Ala Thr Thr
 195 200 205
 Ser Ser Leu Glu Ala Val Pro Phe Ala Ser Val Ala Asn Val Ala Cys
 210 215 220
 Gly Asp Glu Ala Lys Ser Glu Thr His Tyr Phe Leu Cys Asn Glu Lys
 225 230 235 240
 Thr Pro Gly Ile Phe His Trp Gly Ser Ser Gly Pro Leu Cys Val Ser
 245 250 255
 Pro Lys Phe Gly Cys Ser Phe Asn Asn Gly Gly Cys Gln Gln Asp Cys
 260 265 270
 Phe Glu Gly Gly Asp Gly Ser Phe Arg Cys Gly Cys Arg Pro Gly Phe
 275 280 285
 Arg Leu Leu Asp Asp Leu Val Thr Cys Ala Ser Arg Asn Pro Cys Ser
 290 295 300
 Ser Asn Pro Cys Thr Gly Gly Gly Met Cys His Ser Val Pro Leu Ser
 305 310 315 320
 Glu Asn Tyr Thr Cys Arg Cys Pro Ser Gly Tyr Gln Leu Asp Ser Ser
 325 330 335
 Gln Val His Cys Val Asp Ile Asp Glu Cys Gln Asp Ser Pro Cys Ala
 340 345 350
 Gln Asp Cys Val Asn Thr Leu Gly Ser Phe His Cys Glu Cys Trp Val

```

          355          360          365
Gly Tyr Gln Pro Ser Gly Pro Lys Glu Glu Ala Cys Glu Asp Val Asp
 370          375          380
Glu Cys Ala Ala Ala Asn Ser Pro Cys Ala Gln Gly Cys Ile Asn Thr
 385          390          395          400
Asp Gly Ser Phe Tyr Cys Ser Cys Lys Glu Gly Tyr Ile Val Ser Gly
          405          410          415
Glu Asp Ser Thr Gln Cys Glu Asp Ile Asp Glu Cys Ser Asp Ala Arg
          420          425          430
Gly Asn Pro Cys Asp Ser Leu Cys Phe Asn Thr Asp Gly Ser Phe Arg
          435          440          445
Cys Gly Cys Pro Pro Gly Trp Glu Leu Ala Pro Asn Gly Val Phe Cys
          450          455          460
Ser Arg Gly Thr Val Phe Ser Glu Leu Pro Ala Arg Pro Pro Gln Lys
 465          470          475          480
Glu Asp Asn Asp Asp Arg Lys Glu Ser Thr Met Pro Pro Thr Glu Met
          485          490          495
Pro Ser Ser Pro Ser Gly Ser Lys Asp Val Ser Asn Arg Ala Gln Thr
          500          505          510
Thr Gly Leu Phe Val Gln Ser Asp Ile Pro Thr Ala Ser Val Pro Leu
          515          520          525
Glu Ile Glu Ile Pro Ser Glu Val Ser Asp Val Trp Phe Glu Leu Gly
 530          535          540
Thr Tyr Leu Pro Thr Thr Ser Gly His Ser Lys Pro Thr His Glu Asp
 545          550          555          560
Ser Val Ser Ala His Ser Asp Thr Asp Gly Gln Asn Leu Leu Leu Phe
          565          570          575
Tyr Ile Leu Gly Thr Val Val Ala Ile Ser Leu Leu Leu Val Leu Ala
          580          585          590
Leu Gly Ile Leu Ile Tyr His Lys Arg Arg Ala Lys Lys Glu Glu Ile
          595          600          605
Lys Glu Lys Lys Pro Gln Asn Ala Ala Asp Ser Tyr Ser Trp Val Pro
 610          615          620
Glu Arg Ala Glu Ser Gln Ala Pro Glu Asn Gln Tyr Ser Pro Thr Pro
 625          630          635          640
Gly Thr Asp Cys

```

<210> 74
 <211> 27
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Primer

<400> 74
tggtgtcgac gcagagtagc ggccgct 27

<210> 75
<211> 14
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 75
ggcccgggcc ggcc 14

<210> 76
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 76
tcgaggccgc ccgggcc 17

<210> 77
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 77
cagccctcac tccttctc 18

<210> 78
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 78
ggtgggtct ttcattcc 18

<210> 79
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 79
ttcagcaagc cctgactc 18

<210> 80
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 80
gccaccttcg aagcaatc 18

<210> 81
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 81
gagcgggtaca ggagaatg 18

<210> 82
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 82
gcccacccaa ccaaatac 18

<210> 83
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 83
accgcgttct tctgtaac 18

<210> 84
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 84
cagctaacag caggatcc 18

<210> 85
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 85
gacccaggct gcccc 15

<210> 86
<211> 15
<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 86

ggtactcttg ttgag

15

<210> 87

<211> 34

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 87

gaggggtacca agcttcgtgg tccgcaacca gaag

34

<210> 88

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 88

ctcagatctg aattcggcct gtccctcggg gaggag

36

<210> 89

<211> 35

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 89

gacaagatct cagctgaata gcgaccacc ctgtg

35

<210> 90

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 90

gcatctagag cggccgctca ggctgtccc tcggtgagga g

41

<210> 91

<211> 37

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer

<400> 91

gcatctagag cggccgcaa cgctgcttag atctcct

37

<210> 92

<211> 131
 <212> PRT
 <213> Mus musculus

<400> 92
 Glu Val Phe Leu Asn Pro Leu Gly Arg Asp Leu Leu Ser Ile Ser Asp
 1 5 10 15
 Gly Arg Xaa Arg Ala Ile His Arg Arg Gly His Asn Asp Gly Glu Asp
 20 25 30
 Ser Leu Thr His Thr Asp Val Ser Ser Phe Gln Thr Met Asp Gln Met
 35 40 45
 Val Ser Asn Met Arg Asn Tyr Met Gln Lys Leu Glu Arg Asn Thr Gly
 50 55 60
 Gln Leu Ser Val Asp Pro Asn Gly His Ser Phe Cys Ser Ser Ser Val
 65 70 75 80
 Met Thr Tyr Ser Lys Ile Gly Asp Glu Pro Pro Lys Val Phe His Ala
 85 90 95
 Ser Thr Gln Thr Arg Arg Ala Pro Gly Arg Ile Lys Glu Thr Arg Lys
 100 105 110
 Ala Met Arg Asp Ser Asp Ser Gly Thr Arg Lys His Gly Tyr Trp Ser
 115 120 125
 Ser Tyr Pro
 130

<210> 93
 <211> 131
 <212> PRT
 <213> Mus musculus

<400> 93
 Arg Ser Phe Ser Glu Pro Phe Gly Arg Asp Leu Leu Ser Ile Ser Asp
 1 5 10 15
 Gly Arg Gly Arg Ala His Asn Arg Arg Gly His Asn Asp Gly Glu Asp
 20 25 30
 Ser Leu Thr His Thr Asp Val Ser Ser Phe Gln Thr Met Asp Gln Met
 35 40 45
 Val Ser Asn Met Arg Asn Tyr Met Gln Lys Leu Glu Arg Asn Phe Gly
 50 55 60
 Gln Leu Ser Val Asp Pro Asn Gly His Ser Phe Cys Ser Ser Ser Val
 65 70 75 80
 Met Thr Tyr Ser Lys Ile Gly Asp Glu Pro Pro Lys Val Phe Gln Ala
 85 90 95
 Ser Thr Gln Thr Arg Arg Ala Pro Gly Gly Ile Lys Glu Thr Arg Lys
 100 105 110
 Ala Met Arg Asp Ser Asp Ser Gly Leu Glu Lys Met Ala Ile Gly His
 115 120 125
 His Ile His

130

<210> 94
 <211> 159
 <212> PRT
 <213> Mus musculus

<400> 94
 Ala Lys Tyr Val Cys Leu Ala Asn Lys Asn Cys Pro Val Asp Lys Arg
 1 5 10 15
 Arg Arg Asn Arg Cys Gln Tyr Cys Arg Phe Gln Lys Cys Leu Ala Val
 20 25 30
 Gly Met Val Lys Glu Xaa Val Arg Thr Asp Ser Leu Lys Gly Arg Arg
 35 40 45
 Gly Arg Leu Pro Ser Lys Pro Lys Ser Pro Gln Glu Pro Ser Pro Pro
 50 55 60
 Ser Pro Pro Val Ser Leu Ile Ser Ala Leu Val Arg Ala His Val Asp
 65 70 75 80
 Ser Asn Pro Ala Met Thr Ser Leu Asp Tyr Ser Arg Phe Gln Ala Asn
 85 90 95
 Pro Asp Tyr Gln Met Ser Gly Asp Asp Thr Gln His Ile Gln Gln Phe
 100 105 110
 Tyr Asp Leu Leu Thr Gly Ser Met Glu Ile Ile Arg Gly Trp Ala Xaa
 115 120 125
 Xaa Ile Pro Gly Phe Ala Asp Leu Pro Lys Ala Asp Gln Asp Leu Leu
 130 135 140
 Phe Glu Ser Ala Phe Leu Glu Leu Phe Val Leu Arg Leu Ala Tyr
 145 150 155

<210> 95
 <211> 140
 <212> PRT
 <213> Mus musculus

<400> 95
 Ser Gln Thr Arg Gln Gln Gly Pro Leu Arg Ser Ile Met Lys Asp Leu
 1 5 10 15
 His Ser Asp Asp Asn Glu Glu Glu Ser Asp Glu Val Glu Asp Asn Asp
 20 25 30
 Asn Asp Ser Glu Met Glu Arg Pro Val Asn Arg Gly Gly Ser Arg Ser
 35 40 45
 Arg Arg Val Ser Leu Ser Asp Gly Ser Asp Ser Glu Ser Ser Ser Ala
 50 55 60
 Ser Ser Pro Leu His His Glu Pro Pro Pro Pro Leu Leu Lys Thr Asn
 65 70 75 80
 Asn Asn Gln Ile Leu Glu Val Lys Ser Pro Ile Lys Gln Ser Lys Ser
 85 90 95

Asp Lys Gln Ile Ly. sn Gly Glu Cys Asp Lys Ala Tyr Leu Glu
100 105 110

Leu Val Glu Leu His Arg Arg Leu Met Thr Leu Arg Glu Arg His Ile
115 120 125

Leu Gln Gln Ile Val Asn Leu Ile Glu Glu Thr Gly
130 135 140

<210> 96

<211> 652

<212> PRT

<213> Mus musculus

<400> 96

Met Ala Thr Ser Met Gly Leu Leu Leu Leu Leu Leu Leu Leu Thr
1 5 10 15

Gln Pro Gly Ala Gly Thr Gly Ala Asp Thr Glu Ala Val Val Cys Val
20 25 30

Gly Thr Ala Cys Tyr Thr Ala His Ser Gly Lys Leu Ser Ala Ala Glu
35 40 45

Ala Gln Asn His Cys Asn Gln Asn Gly Gly Asn Leu Ala Thr Val Lys
50 55 60

Ser Lys Glu Glu Ala Gln His Val Gln Arg Val Leu Ala Gln Leu Leu
65 70 75 80

Arg Arg Glu Ala Ala Leu Thr Ala Arg Met Ser Lys Phe Trp Ile Gly
85 90 95

Leu Gln Arg Glu Lys Gly Lys Cys Leu Asp Pro Ser Leu Pro Leu Lys
100 105 110

Gly Phe Ser Trp Val Gly Gly Gly Glu Asp Thr Pro Tyr Ser Asn Trp
115 120 125

His Lys Glu Leu Arg Asn Ser Cys Ile Ser Lys Arg Cys Val Ser Leu
130 135 140

Leu Leu Asp Leu Ser Gln Pro Leu Leu Pro Asn Arg Leu Pro Lys Trp
145 150 155 160

Ser Glu Gly Pro Cys Gly Ser Pro Gly Ser Pro Gly Ser Asn Ile Glu
165 170 175

Gly Phe Val Cys Lys Phe Ser Phe Lys Gly Met Cys Arg Pro Leu Ala
180 185 190

Leu Gly Gly Pro Gly Gln Val Thr Tyr Thr Thr Pro Phe Gln Thr Thr
195 200 205

Ser Ser Ser Leu Glu Ala Val Pro Phe Ala Ser Ala Ala Asn Val Ala
210 215 220

Cys Gly Glu Gly Asp Lys Asp Glu Thr Gln Ser His Tyr Phe Leu Cys
225 230 235 240

Lys Glu Lys Ala Pro Asp Val Phe Asp Trp Gly Ser Ser Gly Pro Leu
245 250 255

Cys Val Ser Pro Lys Tyr Gly Cys Asn Phe Asn Asn Gly Gly His
 260 265 270
 Gln Asp Cys Phe Glu Gly Gly Asp Gly Ser Phe Leu Cys Gly Cys Arg
 275 280 285
 Pro Gly Arg Arg Leu Leu Asp Asp Leu Val Thr Cys Ala Ser Arg Asn
 290 295 300
 Pro Cys Ser Ser Ser Pro Cys Arg Gly Gly Ala Thr Cys Val Leu Gly
 305 310 315 320
 Pro His Gly Lys Asn Tyr Thr Cys Arg Cys Pro Gln Gly Tyr Gln Leu
 325 330 335
 Asp Ser Ser Gln Leu Asp Cys Val Asp Val Asp Glu Cys Gln Asp Ser
 340 345 350
 Pro Cys Ala Gln Glu Cys Val Asn Thr Pro Gly Gly Phe Arg Cys Glu
 355 360 365
 Cys Trp Val Gly Tyr Glu Pro Gly Gly Pro Gly Glu Gly Ala Cys Gln
 370 375 380
 Asp Val Asp Glu Cys Ala Leu Gly Arg Ser Pro Cys Ala Gln Gly Cys
 385 390 395 400
 Thr Asn Thr Asp Gly Ser Phe His Cys Ser Cys Glu Glu Gly Tyr Val
 405 410 415
 Leu Ala Gly Glu Asp Gly Thr Gln Cys Gln Asp Val Asp Glu Cys Val
 420 425 430
 Gly Pro Gly Gly Pro Leu Cys Asp Ser Leu Cys Phe Asn Thr Gln Gly
 435 440 445
 Ser Phe His Cys Gly Cys Leu Pro Gly Trp Val Leu Ala Pro Asn Gly
 450 455 460
 Val Ser Cys Thr Met Gly Pro Val Ser Leu Gly Pro Pro Ser Gly Pro
 465 470 475 480
 Pro Asp Glu Glu Asp Lys Gly Glu Lys Glu Gly Ser Thr Val Pro Arg
 485 490 495
 Ala Ala Thr Ala Ser Pro Thr Arg Gly Pro Glu Gly Thr Pro Lys Ala
 500 505 510
 Thr Pro Thr Thr Ser Arg Pro Ser Leu Ser Ser Asp Ala Pro Ile Thr
 515 520 525
 Ser Ala Pro Leu Lys Met Leu Ala Pro Ser Gly Ser Ser Gly Val Trp
 530 535 540
 Arg Glu Pro Ser Ile His His Ala Thr Ala Ala Ser Gly Pro Gln Glu
 545 550 555 560
 Pro Ala Gly Gly Asp Ser Ser Val Ala Thr Gln Asn Asn Asp Gly Thr
 565 570 575
 Asp Gly Gln Lys Leu Leu Leu Phe Tyr Ile Leu Gly Thr Val Val Ala
 580 585 590
 Ile Leu Leu Leu Leu Ala Leu Ala Leu Gly Leu Leu Val Tyr Arg Lys

595								600							605			
Arg	Arg	Ala	Lys	Arg	Glu	Glu	Lys	Lys	Glu	Lys	Lys	Pro	Gln	Asn	Ala			
610						615					620							
Ala	Asp	Ser	Tyr	Ser	Trp	Val	Pro	Glu	Arg	Ala	Glu	Ser	Arg	Ala	Met			
625					630					635					640			
Glu	Asn	Gln	Tyr	Ser	Pro	Thr	Pro	Gly	Thr	Asp	Cys							
				645					650									

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : C12N 15/12, 1/21, 5/10, C07K 14/47, 16/18, A61K 38/17, G01N 33/50, C12Q 1/68	A3	(11) International Publication Number: WO 00/11168 (43) International Publication Date: 2 March 2000 (02.03.00)
(21) International Application Number: PCT/US99/19052 (22) International Filing Date: 20 August 1999 (20.08.99) (30) Priority Data: 09/138,132 21 August 1998 (21.08.98) US (63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application US 09/138,132 (CIP) Filed on 21 August 1998 (21.08.98) (71) Applicant (for all designated States except US): PRINCETON UNIVERSITY [US/US]; New South Building, 5th floor, P.O. Box 36, Princeton, NJ 08544-0036 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): LEMISCHKA, Thor [US/US]; 4 Firestone Court, Princeton, NJ 08540 (US). MOORE, Kateri [US/US]; 248 Hawthorne Street, Princeton, NJ 08540 (US). (74) Agents: YAMIN, Michael, A. et al.; Klauber & Jackson, 411 Hackensack Avenue, Hackensack, NJ 07601 (US).	(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i> (88) Date of publication of the international search report: 22 June 2000 (22.06.00)	
(54) Title: GENES THAT REGULATE HEMATOPOIETIC BLOOD FORMING STEM CELLS AND USES THEREOF (57) Abstract <p>The present invention provides an isolated nucleic acid derived from an isolated hematopoietic stem cell. The present invention additionally provides an isolated hematopoietic stem cell specific protein or a portion thereof encoded by the provided nucleic acid. In addition, the present invention provides a method for generating a stem cell/progenitor cell from a primitive hematopoietic cell. Also, the present invention further provides a method for identifying the presence in a sample of a compound that modulates hematopoietic stem cell activity. The present invention additionally provides a molecularly defined primitive hematopoietic stem cell. Finally, the present invention provides a method of ex vivo expansion of hematopoietic stem cells.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 99/19052

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/12 C12N1/21 C12N5/10 C07K14/47 C07K16/18
 A61K38/17 G01N33/50 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C07K C12Q G01N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>DATABASE GENBANK [Online] Accession No. AA154742, 11 December 1996 (1996-12-11) MARRA M.: "EST; mouse lymph node cDNA clone IMAGE:634660 similar to B-cell activation protein BL34" XP002126589 compare nt 1-390 with nt 13-400 of seq. ID 1.</p> <p>---</p>	<p>1-5, 7-20, 22-31, 42-50</p>
A	<p>HONG J.X. ET AL.: "Isolation and characterization of a novel B-cell activation gene." J. IMMUNOL., vol. 150, no. 9, 1993, pages 3895-3904, XP002126588 the whole document</p> <p>---</p>	<p>1-5, 7-20, 22-31, 42-50</p>
	-/--	



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

* Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

A document member of the same patent family

Date of the actual completion of the international search

22 December 1999

Date of mailing of the international search report

31.03.00

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentaan 2
 NL - 2280 HV Rijswijk
 Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
 Fax: (+31-70) 340-3016

Authorized officer

Galli, I

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 99/19052

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>MATTHEWS W ET AL: "A RECEPTOR TYROSINE KINASE SPECIFIC TO HEMATOPOIETIC STEM AND PROGENITOR CELL-ENRICHED POPULATIONS" CELL,US,CELL PRESS, CAMBRIDGE, NA, vol. 65, no. 7, page 1143-1152 XP000615979 ISSN: 0092-8674 cited in the application -----</p>	<p>1-5, 7-20, 22-31, 42-50</p>

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 99/ 19052

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Remark: Although claims 46-49
are directed to a method of treatment of the human/animal
body, the search has been carried out and based on the alleged
effects of the compound/composition.
2. ☒ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such
an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheets

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all
searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment
of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report
covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is
restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

claims 1-5,7-20,22-31,42-50)- partial

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box 1.2

Claims 31,42,43,46,50 refer to modulators of the polypeptides without giving a true technical characterization. Moreover, no such compounds are defined in the application. In consequence, the scope of said claims is ambiguous and vague, and their subject-matter is not sufficiently disclosed and supported (Art. 5 and 6) PCT. No search can be performed for such purely speculative claims whose wording is, in fact, a mere recitation of the results to be achieved.

Claims 18-19 are unclear with regard to their dependence from claim 16 and to their scope.

Claims 2,10,44 are inconsistent, in that Seq. ID 42 defines an amino acid sequence and not a nucleic acid.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

1. Claims: (1-5,7-20,22-31,42-50) - partial

An isolated nucleic acid derived from an isolated hematopoietic stem cell, expressed in and specific to a hematopoietic stem cell.

Said nucleic acid capable of hybridizing to seq. ID 1.

Corresponding vectors, host systems, polypeptides, antibodies and antibody-producing cells, detection methods, modulators and methods to identify them, therapeutic and pharmecautic applications.

2. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 1.

3. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 3

4. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 4

5. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 5

6. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 6

7. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 7

8. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 8

9. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 9

10. Claims: (1-5,7-20,22-31,42-50) - partial

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Idem as subject matter 1, but limited to seq. ID 10

11. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 11

12. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 12

13. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 13

14. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 14

15. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 15

16. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 16

17. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 17

18. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 18

19. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 19

20. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 20

21. Claims: (1-5,7-20,22-31,42-50) - partial

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Idem as subject matter 1, but limited to seq. ID 21

22. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 22

23. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 23

24. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 24

25. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 25

26. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 26

27. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 27

28. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 28

29. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 29

30. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 30

31. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 31

32. Claims: (1-5,7-20,22-31,42-50) - partial

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Idem as subject matter 1, but limited to seq. ID 32

33. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 33

34. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 34

35. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 35

36. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 36

37. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 37

38. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 38

39. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 39

40. Claims: (1-5,7-20,22-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 40

41. Claims: (1-5,7-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 41 (and for the polypeptide, seq. ID 42)

42. Claims: (1-5,7-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 43 (and for the polypeptide, seq. ID 44)

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

43. Claims: (1-5,7-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 45 (and for the polypeptide, seq. ID 46)

44. Claims: (1-5,7-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 47 (and for the polypeptide, seq. ID 48)

45. Claims: (1-5,7-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 49 (and for the polypeptide, seq. ID 50)

46. Claims: (1-5,7-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 51 (and for the polypeptide, seq. ID 52)

47. Claims: (1-5,7-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 53 (and for the polypeptide, seq. ID 54)

48. Claims: (1-5,7-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 55 (and for the polypeptide, seq. ID 56)

49. Claims: (1-5,7-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 57 (and for the polypeptide, seq. ID 58)

50. Claims: (1-5,7-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 59 (and for the polypeptide, seq. ID 60)

51. Claims: (1-5,7-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 61 (and for the polypeptide, seq. ID 62)

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

52. Claims: (1-5,7-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 63 (and for the polypeptide, seq. ID 64)

53. Claims: (1-5,7-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 65 (and for the polypeptide, seq. ID 66)

54. Claims: (1-5,7-31,42-50) - partial

Idem as subject matter 1, but limited to seq. ID 67 (and for the polypeptide, seq. ID 68)

55. Claims: (17,21) - partial

An isolated hematopoietic stem cell-specific protein or portion thereof.

Said protein characterized by seq. ID 70

56. Claims: (17,21) - partial

Idem as subject matter 56, but limited to seq. ID 71

57. Claims: (1-5,7-20,22-31,42-50) - partial; (6) - complete

Idem as subject matter 1, but limited to seq. ID 72 (and for the polypeptide, seq. ID 73)

58. Claims: (32-41) - complete

A method for identifying primitive hematopoietic stem cell-specific nucleic acids.

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.